# Polynomial optimization

Sun, Zhao

# Polynomial Optimization: Error Analysis and Applications

Tilburg University

Zhao Sun

# Polynomial Optimization: Error Analysis and Applications

Proefschrift

ter verkrijging van de graad van doctor aan
Tilburg University
op gezag van de rector magnificus
prof. dr. E.H.L. Aarts
in het openbaar te verdedigen ten overstaan van een
door het college voor promoties aangewezen commissie
in de aula van de Universiteit
op maandag 29 juni 2015
om 14.15 uur

door

**Zhao Sun**

geboren op 1 december 1987
te Zibo, China

**Promotiecommissie**

| | |
|---|---|
| Promotores: | Prof. Dr. E. de Klerk |
| | Prof. Dr. M. Laurent |
| Overige Leden: | Prof. Dr. D. Henrion |
| | Prof. Dr. R. Sotirov |
| | Prof. Dr. F. Vallentin |
| | Dr. J.C. Vera Lizcano |

# Acknowledgements

This thesis is the outcome of my three-year work as a PhD student at Tilburg University. However, I can never achieve it without the help from a great many people, and I would like to take this opportunity to thank some of them.

First of all, I would like to express my deepest gratitude to my supervisors Monique Laurent and Etienne de Klerk, for their excellent guidance and extremely generous support. Working with them has been an invaluable experience. They were always very patient to answer my questions and teach me the material that I needed to know. When I started to write up my work, they taught me how to write in a legible manner. They also taught me how the academia works and gave me a lot of helpful advice on my future career. I am very impressed with their limitless enthusiasm and the highest standards for mathematical research, and this will be influential in my entire life. In sum, Monique and Etienne provided me with an excellent atmosphere for doing research, and without their help this thesis can never come into existence.

I wish to especially thank the members of my PhD committee, Didier Henrion, Renata Sotirov, Frank Vallentin and Juan Vera, for their helpful comments and suggestions to improve this thesis. I would further like to thank Frank, Juan and Renata for teaching the courses which helped me a lot to develop the background in optimization.

Furthermore, I would like to thank my colleagues and friends from the Operations Research Group for creating a welcoming and hospitable working environment: Aida Abiad, Marleen Balvert, Jac Braat, Ruud Brekelmans, Dick den Hertog, Sybren Huijink, Ning Ma, Krzysztof Postek, Renata Sotirov, Edwin van Dam, Juan Vera and Jianzhe Zhen. In particular, I would like to thank Dick, Edwin and Juan for their help and advice on my job search. Moreover, I would like to thank our sectaries, Korine Bor, Heidi Ket, Lenie Laurijssen and Anja Manders, for their kind support on administrative issues.

Many thanks to my Chinese friends, Xu Lang, Hong Li, Lei Shu, Ruixin Wang, Yifan Yu and Jianzhe Zhen, for the cheerful memories in Tilburg. In particular, I would like to express my gratitude to Jianzhe and Yifan for their constant help in my personal life.

Finally, but most importantly, I would like to thank my mother, my father and my girlfriend Jingwen, for their unconditional full support during the past years. I want to dedicate this thesis to them.

Tilburg, April 2015

# Contents

# Chapter 1

# Introduction

## 1.1 Polynomial optimization

Polynomial optimization, as its name suggests, is the problem of optimizing a polynomial function subject to polynomial inequality constraints. More precisely, given polynomials $f, g_1, \ldots, g_m \in \mathbb{R}[x]$ in $n$ variables $x = (x_1, \ldots, x_n)$, we consider the following optimization problem, which is the general form of a *polynomial optimization problem*:

$$f_{\min,\mathbf{K}} := \inf \; f(x) \;\; \text{s.t. } x \in \mathbf{K} := \{x \in \mathbb{R}^n : g_1(x) \geq 0, \ldots, g_m(x) \geq 0\}. \tag{1.1}$$

Analogously, we denote

$$f_{\max,\mathbf{K}} := \sup \; f(x) \;\; \text{s.t. } x \in \mathbf{K} = \{x \in \mathbb{R}^n : g_1(x) \geq 0, \ldots, g_m(x) \geq 0\}.$$

In particular, when the polynomials $f, g_1, \ldots, g_m$ are affine, problem (1.1) becomes a linear programming problem. Thus, polynomial optimization contains linear programming (LP) as a special case. Moreover, since the binary integrality constraints $x_i \in \{0, 1\}$ ($i \in [n] := \{1, \ldots, n\}$) can be expressed by the polynomial equality constraints $x_i(1 - x_i) = 0$ ($i \in [n]$), polynomial optimization also captures 0-1 integer linear programming, where the constraints $x_i \in \{0, 1\}$ ($i \in [n]$) are added to general linear programs.

In this thesis, we will mainly consider the special cases where the set $\mathbf{K}$ is a standard simplex or a hypercube. For an overview of polynomial optimization, see, e.g., the book [3] by Anjos and Lasserre.

### 1.1.1 Applications

Polynomial optimization is a fundamental model in optimization and has very wide applications, e.g., in combinatorial optimization, control theory, signal processing and mathematical finance. To motivate our study, we illustrate some sample applications of polynomial optimization below.

In fact, many combinatorial optimization problems can be formulated as 0-1 integer linear programs; this is the case, e.g., for assignment, scheduling and packing problems (see, e.g., [88]). Thus, they can be reformulated as polynomial optimization problems, since polynomial optimization contains 0-1 integer linear programming as a special case. In particular, we recall two hard problems in graphs, the maximum stable set problem and the maximum cut (max-cut) problem, which we will consider in this thesis. As we see below, they can be reformulated as polynomial optimization over the standard simplex and the hypercube, respectively.

Given a graph $G = (V, E)$, a set $S \subseteq V$ is stable if no two distinct nodes of $S$ are adjacent in $G$. The *maximum stable set problem* asks to find the maximum cardinality of a stable set in $G$, which is denoted by $\alpha(G)$ and called the *stability number* of $G$. Let $A$ denote the adjacency matrix of $G$ and let $I$ denote the identity matrix. Then, by a result of Motzkin and Straus [70], the stability number $\alpha(G)$ can be obtained via

$$\frac{1}{\alpha(G)} = \min_{x \in \Delta_{|V|}} x^T(I + A)x, \tag{1.2}$$

where

$$\Delta_{|V|} := \left\{ x \in \mathbb{R}_+^{|V|} : \sum_{i=1}^{|V|} x_i = 1 \right\} \tag{1.3}$$

denotes the standard simplex.

Given a graph $G = (V, E)$ with edge weights $w \in \mathbb{R}^{|E|}$, the *max-cut problem* asks to find a partition $(V_1, V_2)$ of the node set $V$ so that the total weight of the edges cut by the partition is maximized. As observed in [77], setting $d_i = \sum_{j \in V : ij \in E} w_{ij}$, the maximum weight of a cut in $G$, denoted as $\mathrm{mc}(G, w)$, can be computed via the following polynomial optimization problem:

$$\mathrm{mc}(G, w) = \max_{x \in [0,1]^{|V|}} \sum_{i \in V} d_i x_i - 2 \sum_{ij \in E} w_{ij} x_i x_j. \tag{1.4}$$

In mathematical finance, a representative instance would be the so-called mean-variance-skewness-kurtosis portfolio decision problem. Given $n$ risky assets, we denote the return of asset $i \in [n]$ at the end of a planning period by $R_i$ (seen as a random variable). A portfolio is represented by a point $x \in \Delta_n$, where $x_i$ ($i \in [n]$) denotes the proportion of the investor's capital invested in asset $i$. Thus the return on the portfolio is the random variable $R = \sum_{i=1}^n x_i R_i$, say. Let $\mu_i = \mathbb{E}[R_i]$ ($i \in [n]$) so that the expected return on the portfolio is $\mathbb{E}[R] = \sum_{i=1}^n x_i \mu_i$. Similarly, for $i, j, k, l \in [n]$, let

$$\begin{aligned}
\sigma_{ij} &= \mathbb{E}[(R_i - \mu_i)(R_j - \mu_j)], \\
\varsigma_{ijk} &= \mathbb{E}[(R_i - \mu_i)(R_j - \mu_j)(R_k - \mu_k)], \\
\kappa_{ijkl} &= \mathbb{E}[(R_i - \mu_i)(R_j - \mu_j)(R_k - \mu_k)(R_l - \mu_l)].
\end{aligned}$$

(In practice these values are estimated from historical data.) Now the variance of $R$ is $\mathbb{E}[(R - \mathbb{E}(R))^2] = \sum_{i,j=1}^n x_i x_j \sigma_{ij}$; the skewness of $R$ is $\mathbb{E}[(R - \mathbb{E}(R))^3] = \sum_{i,j,k=1}^n x_i x_j x_k \varsigma_{ijk}$; the kurtosis of $R$ is $\mathbb{E}[(R - \mathbb{E}(R))^4] = \sum_{i,j,k,l=1}^n \kappa_{ijkl} x_i x_j x_k x_l$. The goal is to maximize the expected value of $R$ as well as its skewness, while minimizing the variance and kurtosis (seen as risk measures). The portfolio decision problem then becomes

$$\begin{aligned}
\max \quad & \lambda_1 \sum_{i=1}^n \mu_i x_i - \lambda_2 \sum_{i,j=1}^n \sigma_{ij} x_i x_j + \lambda_3 \sum_{i,j,k=1}^n \varsigma_{ijk} x_i x_j x_k - \lambda_4 \sum_{i,j,k,l=1}^n \kappa_{ijkl} x_i x_j x_k x_l \\
\text{s.t.} \quad & x \in \Delta_n,
\end{aligned}$$

where the nonnegative parameters $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ measure the investor's preference to the four moments, and they sum up to one, i.e., $\sum_{i=1}^4 \lambda_i = 1$; see, e.g., [42, 43] for more details on the model. In particular, if one only considers the first two central moments (i.e., $\lambda_3 = \lambda_4 = 0$), the above model becomes the celebrated Markowitz's mean-variance model [66]. For more applications in mathematical finance, see, e.g., [52] and the references therein.

### 1.1.2 Relaxation methods for polynomial optimization

From a complexity point of view, while there exist polynomial time algorithms for solving linear programming, no such algorithm is known for general polynomial optimization. The polynomial optimization problem (1.1) is in fact NP-hard in general, even for some special cases. In fact, minimizing a quadratic polynomial over the standard simplex or the unit hypercube is already NP-hard, as it contains the maximum stable set problem in (1.2) and the max-cut problem in (1.4) as special cases, and these two problems are NP-hard by Karp [41]. For NP-hard optimization problems, a common strategy is to design some tractable relaxations, which can give upper/lower bounds for the optimal value in polynomial time. For more information about the complexity and relaxation methods for polynomial optimization, see, e.g., the book [61] by Li et al. and the survey [16] by De Klerk.

In particular, about fifteen years ago, Lasserre [47] and Parrilo [79, 80] proposed the so-called *SOS* (sums of squares) method for the polynomial optimization problem. It uses sums of squares of polynomials to construct tractable hierarchies of approximations, which converge asymptotically to the global optimum. This method is based on some celebrated developments in real algebraic geometry which give representations for positive polynomials in terms of sums of squares of polynomials. It also uses the fact that deciding whether a polynomial is a sum of squares of polynomials can be verified via a semidefinite program. Indeed, testing whether a polynomial $\sigma$ of degree $2r$ is a sum of squares of polynomials amounts to testing whether there exists a positive semidefinite matrix $M$ of order $\binom{n+r}{r}$, satisfying $\sigma(x) = [x]_r^T M [x]_r$, where $[x]_r = (x^\alpha)_{\alpha\in\mathbb{N}^n,\sum_{i=1}^n \alpha_i \le r}$ is the vector containing all monomials of degree at most $r$. Then by equating the coefficients of the monomials in the polynomials $\sigma(x)$ and $[x]_r^T M[x]_r$, we find a semidefinite program involving a matrix of size $\binom{n+r}{r}$. Recall that semidefinite programming (SDP) is a generalization of linear programming, where vector variables are replaced by positive semidefinite matrix variables (see, e.g., [56] for an overview). In recent years, semidefinite programming has been used in many relaxation methods, since it can be solved in polynomial time to any fixed accuracy, e.g., by the ellipsoid method [6, 91] or the interior point method [85, 15].

When applying the SOS method to problem (1.1), one starts with reformulating $f_{\min,\mathbf{K}}$ as

$$f_{\min,\mathbf{K}} = \sup\ \lambda \text{ s.t. } f - \lambda \text{ is nonnegative on } \mathbf{K}.$$

Recall that a polynomial $f$ is nonnegative (resp., positive) on $\mathbf{K}$ if $f(x) \ge 0$ (resp., $f(x) > 0$) for all $x \in \mathbf{K}$.

Then, lower bounds for $f_{\min,\mathbf{K}}$ can be obtained by using sufficient conditions to replace the nonnegativity of $f - \lambda$ on $\mathbf{K}$. These representations lead to hierarchies of relaxations that can be computed with linear programming or semidefinite programming. See Section 1.2.1 for more details.

Additionally, there is another type of approach proposed by Lasserre [47, 53] to construct hierarchies of upper bounds for the minimum $f_{\min,\mathbf{K}}$. The idea is to reformulate the problem of computing $f_{\min,\mathbf{K}}$ as the problem of finding a probability measure with support $\mathbf{K}$ such that the expected value of $f$ under this measure is minimized. Then, by selecting some probability measure on $\mathbf{K}$, one can obtain an upper bound for $f_{\min,\mathbf{K}}$. For more details, see Section 1.2.2.

In this thesis, we study several popular hierarchies of relaxations. Our main interest lies in understanding their performance, in particular how fast they converge to the global optimum. Then, by getting good estimates on the rate of convergence, one can judge the qualities of these hierarchies. Next we introduce these hierarchies of relaxations.

# 1.2 Hierarchies of relaxations

## 1.2.1 Representations for positive polynomials

In this section we introduce some approaches to construct hierarchies of lower bounds for $f_{\min,\mathbf{K}}$ as already mentioned before. With $\mathscr{P}(\mathbf{K})$ denoting the set of real polynomials that are nonnegative on the set $\mathbf{K}$, problem (1.1) can be rewritten as

$$f_{\min,\mathbf{K}} = \sup\ \lambda \text{ s.t. } f-\lambda \in \mathscr{P}(\mathbf{K}). \tag{1.5}$$

In the above formula, the nonnegativity condition $f-\lambda \in \mathscr{P}(\mathbf{K})$ is hard to test in general. Then, the idea is to replace the hard nonnegativity condition by some tractable and sufficient conditions. For instance, if the polynomial $f-\lambda$ can be written as $f-\lambda = \sum_{\alpha\in\mathbb{N}^m} c_\alpha g_1^{\alpha_1}\cdots g_m^{\alpha_m}$, where the parameters $c_\alpha$ are nonnegative scalars or more generally sums of squares of polynomials, then $f-\lambda$ must be nonnegative on the set $\mathbf{K} = \{x\in\mathbb{R}^n : g_1(x)\geq 0,\ldots,g_m(x)\geq 0\}$. Based on these conditions, one can construct LP/SDP-based hierarchies of lower bounds for $f_{\min,\mathbf{K}}$.

In what follows, we introduce four results, by Pólya, Handelman, Schmüdgen and Putinar, respectively, which give different types of representations for positive polynomials. We recommend the references [67, 52, 58] for an overview. Among these results, the results by Pólya and Handelman lead to LP-based hierarchies of lower bounds, while the results of Schmüdgen and Putinar lead to SDP-based approximations.

Throughout this section, we consider a polynomial $f$ of degree $d$, written as $f = \sum_{\beta\in N(n,d)} f_\beta x^\beta$, where $x^\beta := \prod_{i=1}^n x_i^{\beta_i}$ and

$$N(n,d) := \{\beta\in\mathbb{N}^n : \sum_{i=1}^n \beta_i \leq d\}.$$

Recall that the degree of the monomial $x^\beta$ is $|\beta| := \sum_{i=1}^n \beta_i$ and the degree of the polynomial $f = \sum_{\beta\in\mathbb{N}^n} f_\beta x^\beta$ is defined as $\deg(f) := \max_{\{\beta: f_\beta\neq 0\}} |\beta|$. We also consider the parameter

$$L(f) := \max_{\beta} \frac{\beta!}{|\beta|!}|f_\beta|, \tag{1.6}$$

where $\beta! := \beta_1! \cdots \beta_n!$.

### Pólya's representation theorem

We first consider the special case when the set $\mathbf{K}$ is the standard simplex $\Delta_n = \left\{x \in \mathbb{R}^n_+ : \sum_{i=1}^n x_i = 1\right\}$ from (1.3).

This is already interesting, since the problem of computing $f_{\min,\Delta_n}$, the minimum of $f$ on $\Delta_n$, contains the maximum stable set problem (1.2) as a special case. Note that one can assume w.l.o.g. that $f$ is homogeneous, which means that all monomials in $f$ have the same degree.

One can easily see that if the polynomial $\left(\sum_{i=1}^n x_i\right)^r f(x)$ has nonnegative coefficients for some integer $r \geq 1$, then $f$ must be nonnegative on $\Delta_n$. In fact, Pólya [82] proves that the reverse implication also holds if we restrict to polynomials that are strictly positive on $\Delta_n$. Moreover, Powers and Reznick [83] give an explicit bound on the degree $r$ for which $\left(\sum_{i=1}^n x_i\right)^r f(x)$ has nonnegative coefficients.

**Theorem 1.1.** *[82, 83] Suppose $f$ is a homogeneous polynomial of degree $d$ and consider the parameter $L(f)$ from (1.6). If $f$ is positive on the standard simplex $\Delta_n$, then the polynomial $\left(\sum_{i=1}^n x_i\right)^r f(x)$ has nonnegative coefficients for all $r$ satisfying*

$$r \geq \binom{d}{2} \frac{L(f)}{f_{\min,\Delta_n}} - d.$$

Since $f$ is homogeneous of degree $d$, $f_{\min,\Delta_n}$ can be equivalently formulated as

$$f_{\min,\Delta_n} = \sup \quad \lambda \ \text{ s.t. } \ f(x) - \lambda \left(\sum_{i=1}^n x_i\right)^d \geq 0 \ \ \forall x \in \mathbb{R}^n_+. \tag{1.7}$$

Indeed, from (1.5), we have $f_{\min,\Delta_n} = \sup\{\lambda : f(x) - \lambda \geq 0, \forall x \in \Delta_n\}$. Note that $f(x) - \lambda \geq 0$ for any $x$ in $\Delta_n$ if and only if $f\left(y/(\sum_{i=1}^n y_i)\right) - \lambda \geq 0$ for any nonzero $y$ in $\mathbb{R}^n_+$. Then, combining with the fact that $f$ is homogeneous of degree $d$, we obtain the above formulation (1.7).

Then, based on Theorem 1.1, a lower bound for $f_{\min,\Delta_n}$ can be constructed as follows. For any integer $r \geq d$, define the parameter

$$f_{\min}^{(r-d)} := \sup \lambda \text{ s.t. } \left(\sum_{i=1}^{n} x_i\right)^{r-d} \left(f - \lambda \left(\sum_{i=1}^{n} x_i\right)^{d}\right) \in \mathbb{R}_+[x], \tag{1.8}$$

where $\mathbb{R}_+[x]$ denotes the set of polynomials with nonnegative coefficients.

Observe that the parameters $f_{\min}^{(r-d)}$ with increasing $r$ form a hierarchy of lower bounds for $f_{\min,\Delta_n}$, i.e.,

$$f_{\min}^{(0)} \leq f_{\min}^{(1)} \leq \cdots \leq f_{\min}^{(r)} \leq \cdots \leq f_{\min,\Delta_n}.$$

Note that, for fixed $r \geq d$, the parameter $f_{\min}^{(r-d)}$ can be computed via a linear program in the variable $\lambda$. This linear program is obtained by checking the nonnegativity for the coefficients of the monomials $x^\alpha$ for $\alpha \in \mathbb{N}^n$ in the polynomial

$$\left(\sum_{i=1}^{n} x_i\right)^{r-d} \left(f - \lambda \left(\sum_{i=1}^{n} x_i\right)^{d}\right).$$

Based on this, for any polynomial $f = \sum_{\beta \in \mathbb{N}^n} f_\beta x^\beta$ of degree $d$, one can prove (see Lemma 4.1 for details) that

$$f_{\min}^{(r-d)} = \min_{\alpha \in I(n,r)} \sum_{\beta \in \mathbb{N}^n} f_\beta \frac{\alpha^{\underline{\beta}}}{r^{\underline{d}}},$$

where $I(n,r) := \{x \in \mathbb{N}^n : \sum_{i=1}^n x_i = r\}$, $r^{\underline{d}} := r(r-1)\cdots(r-d+1)$ and $\alpha^{\underline{\beta}} := \prod_{i=1}^n \alpha_i^{\underline{\beta_i}}$ for $\alpha, \beta \in \mathbb{N}^n$. Thus, one can compute $f_{\min}^{(r-d)}$ by $|I(n,r)| = \binom{n+r-1}{r}$ evaluations of the polynomial $\sum_{\beta \in \mathbb{N}^n} f_\beta \frac{x^{\underline{\beta}}}{r^{\underline{d}}}$ at the points $x \in I(n,r)$.

For more information on the hierarchical approximations based on Pólya's representation theorem, see, e.g., [14, 21, 92]. In particular, De Klerk et al. [21] study the quality of the bounds $f_{\min}^{(r-d)}$ and show the following upper estimates for $f_{\min,\Delta_n} - f_{\min}^{(r-d)}$ in terms of $f_{\max,\Delta_n} - f_{\min,\Delta_n}$, the range of values of $f$ on $\Delta_n$.

**Theorem 1.2.** *(i) [21, Theorem 1.3] Let $f$ be a homogeneous quadratic polynomial and $r \geq 2$ an integer. Then, one has*

$$f_{\min,\Delta_n} - f_{\min}^{(r-2)} \leq \frac{1}{r-1}(f_{\max,\Delta_n} - f_{\min,\Delta_n}).$$

*(ii) [21, Theorem 3.2] Let $f$ be a homogenous polynomial of degree $d$ and $r \geq d$ an integer. Then, one has*

$$f_{\min,\Delta_n} - f_{\min}^{(r-d)} \leq \left(\frac{r^d}{r^{\underline{d}}} - 1\right) \frac{2d-1}{d} d^d (f_{\max,\Delta_n} - f_{\min,\Delta_n}).$$

Later in Chapter 4, we will consider the lower bound $f_{\min}^{(r-d)}$ together with the following upper bound $f_{\Delta(n,r)}$ for $f_{\min,\Delta_n}$, defined as

$$f_{\Delta(n,r)} := \min \ f(x) \text{ s.t. } x \in \Delta(n,r) := \{x \in \Delta_n : rx \in \mathbb{N}^n\}. \tag{1.9}$$

(For more details about $f_{\Delta(n,r)}$, see Chapters 2, 3 and 4.) More precisely, we will study the link between the two parameters $f_{\Delta(n,r)}$ and $f_{\min}^{(r-d)}$. This will enable us to prove upper bounds for the range $f_{\Delta(n,r)} - f_{\min}^{(r-d)}$ that refine earlier results obtained by separately upper bounding the two ranges $f_{\Delta(n,r)} - f_{\min,\Delta_n}$ and $f_{\min,\Delta_n} - f_{\min}^{(r-d)}$. See Chapter 4 for more details.

### Handelman's representation theorem

When the set $\mathbf{K}$ is a full-dimensional polytope, Handelman [38] shows the following result.

**Theorem 1.3.** *[38] Assume that the set $\mathbf{K} = \{x \in \mathbb{R}^n : g_1(x) \geq 0, \ldots, g_m(x) \geq 0\}$ in (1.1) is a full-dimensional polytope and that its defining polynomials $g_1, \ldots, g_m$ are linear polynomials. For any polynomial $f \in \mathbb{R}[x]$, if $f$ is strictly positive on $\mathbf{K}$, then it can be written as*

$$f = \sum_{\alpha \in \mathbb{N}^m} c_\alpha g_1^{\alpha_1} \cdots g_m^{\alpha_m}, \quad \textit{for scalars } c_\alpha \geq 0, \tag{1.10}$$

*where $c_\alpha > 0$ holds for finitely many $\alpha \in \mathbb{N}^m$.*

Powers and Reznick [83] give a constructive proof for Theorem 1.3 and give an upper bound for the degree of the polynomials involved in the representation (1.10). Moreover, a more general result holds when $\mathbf{K}$ is a compact semialgebraic set, as proved by Krivine [45, 46], see, e.g., [58] and the references therein.

We now present a hierarchy of lower bounds for $f_{\min,\mathbf{K}}$ based on Theorem 1.3. We let $g$ denote the set of polynomials $g_1, \ldots, g_m$. For an integer $r \geq 1$, define the *Handelman set of order $r$* as

$$\mathcal{H}_r(g) := \left\{ \sum_{\alpha \in N(m,r)} c_\alpha g_1^{\alpha_1} \cdots g_m^{\alpha_m} \ : \ c_\alpha \geq 0 \text{ for all } \alpha \in N(m,r) \right\},$$

and the corresponding *Handelman bound of order* $r$ as

$$f_{\mathrm{han}}^{(r)} := \sup\{\lambda : f - \lambda \in \mathcal{H}_r(g)\}. \tag{1.11}$$

Clearly, any polynomial in $\mathcal{H}_r(g)$ is nonnegative on $\mathbf{K}$ and one has the following chain of inclusions:

$$\mathcal{H}_1(g) \subseteq \ldots \subseteq \mathcal{H}_r(g) \subseteq \mathcal{H}_{r+1}(g) \subseteq \ldots \subseteq \mathscr{P}(\mathbf{K}),$$

giving the chain of inequalities:

$$f_{\min,\mathbf{K}} \geq \cdots \geq f_{\mathrm{han}}^{(r+1)} \geq f_{\mathrm{han}}^{(r)} \geq \cdots \geq f_{\mathrm{han}}^{(1)} \text{ for } r \geq 1.$$

When $\mathbf{K}$ is a full-dimensional polytope and $g_1, \ldots, g_m$ are linear polynomials, the asymptotic convergence of the bounds $f_{\mathrm{han}}^{(r)}$ to $f_{\min,\mathbf{K}}$ (as the order $r$ increases) is guaranteed by Theorem 1.3 above.

Moreover, for fixed $r$, $f_{\mathrm{han}}^{(r)}$ can be computed via a linear program in the variables $c_\alpha$, obtained by identifying the coefficients of the monomials on both sides of the equality $f - \lambda = \sum_{\alpha \in N(m,r)} c_\alpha g_1^{\alpha_1} \cdots g_m^{\alpha_m}$.

We mention two cases where some results are known about the quality of the Handelman bounds, when $\mathbf{K}$ is the standard simplex or the hypercube. These two specific cases are already interesting to study, since they capture some well-known NP-hard problems, e.g., the maximum stable set problem (1.2) and the max-cut problem (1.4).

**Application to optimization on the standard simplex.** We first consider the case when $\mathbf{K}$ in (1.1) is the standard simplex $\Delta_n$, which can be written as

$$\Delta_n = \left\{ x \in \mathbb{R}^n : x_i \geq 0 \ (i \in [n]), 1 - \sum_{i=1}^{n} x_i \geq 0, \sum_{i=1}^{n} x_i - 1 \geq 0 \right\}. \tag{1.12}$$

It turns out that the corresponding Handelman bound $f_{\mathrm{han}}^{(r)}$ coincides with the LP bound $f_{\min}^{(r-d)}$ introduced in (1.8), as proved in the following Lemma 1.4. Therefore, the results of De Klerk et al. [21] in Theorem 1.2 for $f_{\min}^{(r-d)}$ also hold for $f_{\mathrm{han}}^{(r)}$.

**Lemma 1.4.** *Let $f$ be a homogeneous polynomial of degree $d$. Consider the bound $f_{\mathrm{han}}^{(r)}$ from* (1.11) *defined for the standard simplex (in* (1.12)*) and the parameter $f_{\min}^{(r-d)}$ defined in* (1.8)*. For any integer $r \geq d$, one has*

$$f_{\mathrm{han}}^{(r)} = f_{\min}^{(r-d)}.$$

**Proof.** The proof is similar as the proof of [20, Proposition 2], and we give it here for clarity. Let $\langle 1-\sum_{i=1}^n x_i\rangle$ denote the ideal in $\mathbb{R}[x]$ generated by the polynomial $1-\sum_{i=1}^n x_i$ and, for an integer $r$, let $\langle 1-\sum_{i=1}^n x_i\rangle_r$ denote its truncation at degree $r$, consisting of all polynomials of the form $u(1-\sum_{i=1}^n x_i)$ where $u\in\mathbb{R}[x]$ has degree at most $r-1$. Moreover, let $\mathbb{R}_+[x]_r$ be the subset of $\mathbb{R}_+[x]$ consisting of polynomials of degree at most $r$. With $g$ standing for the set of polynomials

$$\left\{x_1,\ldots,x_n,\pm\left(1-\sum_{i=1}^n x_i\right)\right\},$$

one can easily see that the Handelman set of order $r$ is given by

$$\mathcal{H}_r(g)=\mathbb{R}_+[x]_r+\langle 1-\sum_{i=1}^n x_i\rangle_r.$$

Assume first that $(f-\lambda\left(\sum_{i=1}^n x_i\right)^d)\left(\sum_{i=1}^n x_i\right)^{r-d}\in\mathbb{R}_+[x]$ for some scalar $\lambda\in\mathbb{R}$. By writing $\sum_{i=1}^n x_i=1+(\sum_{i=1}^n x_i-1)$ and expanding the products $(\sum_{i=1}^n x_i)^d$ and $(\sum_{i=1}^n x_i)^r$, one obtains a decomposition of $f-\lambda$ in $\mathbb{R}_+[x]_r+\langle 1-\sum_{i=1}^n x_i\rangle_r$. This shows the inequality $f_{\text{han}}^{(r)}\le f_{\min}^{(r-d)}$.
Conversely, assume that $f-\lambda\in\mathbb{R}_+[x]_r+\langle 1-\sum_{i=1}^n x_i\rangle_r$ for some scalar $\lambda\in\mathbb{R}$. This implies that $f-\lambda(\sum_{i=1}^n x_i)^d=q+u(1-\sum_{i=1}^n x_i)$, where $q\in\mathbb{R}_+[x]_{r+d}$ and $u\in\mathbb{R}[x]_{r+d-1}$. By evaluating both sides at $\frac{x}{\sum_{i=1}^n x_i}$ and multiplying throughout by $(\sum_{i=1}^n x_i)^r$, we obtain that

$$\left(\sum_{i=1}^n x_i\right)^{r-d}\left(f-\lambda\left(\sum_{i=1}^n x_i\right)^d\right)=q\left(\frac{x}{\sum_{i=1}^n x_i}\right)\left(\sum_{i=1}^n x_i\right)^r\in\mathbb{R}_+[x],$$

since $q$ has degree at most $r$. This shows the reverse inequality $f_{\text{han}}^{(r)}\ge f_{\min}^{(r-d)}$. $\square$

**Application to optimization on the hypercube.** We now turn to the case when $\mathbf{K}$ is the hypercube $\mathbf{Q}_n:=[0,1]^n$. Using Bernstein approximations, De Klerk and Laurent [18] show the following error estimates for the Handelman hierarchy.

**Theorem 1.5.** *[18, Theorem 1.4] Let $\mathbf{K}=\mathbf{Q}_n=[0,1]^n$ and let $g$ stand for the set of polynomials $x_1,\ldots,x_n,1-x_1,\ldots,1-x_n$. Recall that the parameter $L(f)$ is defined in* (1.6). *When $f$ is a polynomial of degree $d$, we have:*
*(i) If $f$ is positive on $\mathbf{K}$, then $f\in\mathcal{H}_r(g)$ for some integer*

$$r\le n\left\lceil\frac{L(f)}{f_{\min,\mathbf{Q}_n}}\binom{d+1}{3}n^d\right\rceil.$$

*(ii) For any integer $t \geq 1$, we have*

$$f - f_{\min,\mathbf{Q}_n} + \frac{L(f)}{t}\binom{d+1}{3}n^d \in \mathcal{H}_r(g) \text{ for some integer } r \leq \max\{tn, d\}.$$

*(iii) For any integer $t \geq d$, we have*

$$f_{\min,\mathbf{Q}_n} - f_{\text{han}}^{(tn)} \leq \frac{L(f)}{t}\binom{d+1}{3}n^d.$$

In the quadratic case a better estimate can be shown.

**Theorem 1.6.** *[18, Theorem 2.1] Let $f = x^T Ax + b^T x$ be a quadratic polynomial. For any integer $t \geq 1$,*

$$f_{\min,\mathbf{Q}_n} - f_{\text{han}}^{(tn)} \leq \frac{\sum_{i:A_{ii}>0} A_{ii}}{t}.$$

We observe that the above result in Theorem 1.6 holds only for relaxations $f_{\text{han}}^{(r)}$ with order $r \geq n$. Moreover, if $f$ is a square-free quadratic polynomial (i.e., $A_{ii} = 0$ for all $i$), then the equality $f_{\min,\mathbf{Q}_n} = f_{\text{han}}^{(n)}$ holds and the Handelman relaxation of order $n$ gives the exact value $f_{\min,\mathbf{Q}_n}$.

For order $r \leq n$, Park and Hong [78] give an error analysis in the quadratic square-free case (see Theorem 6.7). This error analysis applies in particular to the bounds obtained by applying the Handelman hierarchy to the maximum stable set problem. Indeed, the maximum stable set problem can also be reformulated as a square-free quadratic polynomial optimization problem over the hypercube (see (1.20) below). This motivates us to investigate Handelman's hierarchy for the maximum stable set problem. Chapter 6 is devoted to this issue.

### Schmüdgen's Positivstellensatz

Recall that Pólya's theorem holds when $\mathbf{K}$ is the standard simplex, while Handelman's theorem holds when $\mathbf{K}$ is a polytope, and both of them lead to LP-based hierarchies of lower bounds for $f_{\min,\mathbf{K}}$. Now we consider *Schmüdgen's Positivstellensatz* [87], which holds when $\mathbf{K}$ is a general compact set, and leads to an SDP-based hierarchy of lower bounds for $f_{\min,\mathbf{K}}$.

**Theorem 1.7.** *[87] Assume the set $\mathbf{K}$ in (1.1) is compact. For any polynomial $f \in \mathbb{R}[x]$, if $f$ is strictly positive on $\mathbf{K}$, then $f$ can be written as*

$$f = \sum_{\alpha \in \{0,1\}^m} \sigma_\alpha g_1^{\alpha_1} \cdots g_m^{\alpha_m}, \quad \text{where } \sigma_\alpha \text{ are sums of squares of polynomials.} \tag{1.13}$$

We let $\Sigma[x]$ be the set of sums of squares of polynomials. Then, for an integer $r \geq 1$, define the *truncated pre-ordering* as

$$\mathcal{T}_r(g) := \left\{ \sum_{\alpha \in \{0,1\}^m : |\alpha| \leq r} \sigma_\alpha g_1^{\alpha_1} \cdots g_m^{\alpha_m} \ : \ \deg\{\sigma_\alpha g_1^{\alpha_1} \cdots g_m^{\alpha_m}\} \leq r, \ \sigma_\alpha \in \Sigma[x] \right\}$$

and the corresponding *Schmüdgen bound of order $r$* as

$$f_{sch}^{(r)} := \sup\{\lambda : f - \lambda \in \mathcal{T}_r(g)\}.$$

Similarly as for $\mathcal{H}_r(g)$ and $f_{\mathrm{han}}^{(r)}$, one has

$$\mathcal{T}_1(g) \subseteq \ldots \subseteq \mathcal{T}_r(g) \subseteq \mathcal{T}_{r+1}(g) \subseteq \ldots \subseteq \mathscr{P}(\mathbf{K}),$$

giving the chain of inequalities:

$$f_{\min,\mathbf{K}} \geq \cdots \geq f_{sch}^{(r+1)} \geq f_{sch}^{(r)} \geq \cdots \geq f_{sch}^{(1)} \text{ for } r \geq 1.$$

The asymptotic convergence of the bounds $f_{sch}^{(r)}$ to $f_{\min,\mathbf{K}}$ (as $r$ increases) follows directly from Theorem 1.7.

For fixed $r$, the bound $f_{sch}^{(r)}$ can be computed via a semidefinite program. Recall that checking whether a polynomial is a sum of squares of polynomials can be expressed as a semidefinite program. Hence, the problem of testing membership in $\mathcal{T}_r(g)$ can be reformulated as a semidefinite program involving $2^m$ positive semidefinite matrices of order at most $\binom{n+\lfloor r/2 \rfloor}{\lfloor r/2 \rfloor}$.

In addition, one can easily see that $\mathcal{H}_r(g) \subseteq \mathcal{T}_r(g)$. Then, for any integer $r \geq 1$,

$$f_{\mathrm{han}}^{(r)} \leq f_{sch}^{(r)} \leq f_{\min,\mathbf{K}}$$

holds. Thus, if $\mathbf{K} = [0,1]^n$, then the results for the parameter $f_{\mathrm{han}}^{(r)}$ in Theorems 1.5 and 1.6 also hold for the parameter $f_{sch}^{(r)}$. Moreover, by Lemma 1.4, if $\mathbf{K} = \Delta_n$, then

$$f_{\mathrm{han}}^{(r)} = f_{\min}^{(r-d)} \leq f_{sch}^{(r)} \leq f_{\min,\mathbf{K}},$$

and thus the results in Theorem 1.2 for the parameter $f_{\min}^{(r-d)}$ also hold for the parameter $f_{sch}^{(r)}$.

In the general case, when $\mathbf{K}$ is contained in $(-1,1)^n$, Schweighofer [89] gives the following error analysis for the parameter $f_{sch}^{(r)}$.

**Theorem 1.8.** *[89] Assume the set* $\mathbf{K}$ *in* (1.1) *satisfies* $\mathbf{K} \subseteq (-1,1)^n$ *and consider the parameter* $L(f)$ *from* (1.6). *Then, there exist integers* $c, c' > 0$ *satisfying the following properties:*

*(i) Every polynomial* $f$ *of degree* $d$ *which is positive on* $\mathbf{K}$ *belongs to* $\mathcal{T}_r(g)$ *for some integer* $r$ *satisfying*

$$r \leq cd^2 \left[1 + \left(d^2 n^d \frac{L(f)}{f_{\min,\mathbf{K}}}\right)^c\right].$$

*(ii) For every polynomial* $f$ *of degree* $d$ *and for all integers* $r \geq c' d^{c'} n^{c'd}$, *we have*

$$f - f_{\min,\mathbf{K}} + \frac{c' d^4 n^{2d}}{\sqrt[c']{r}} L(f) \in \mathcal{T}_r(g), \text{ and thus } f_{\min,\mathbf{K}} - f_{sch}^{(r)} \leq \frac{c' d^4 n^{2d}}{\sqrt[c']{r}} L(f).$$

### Putinar's Positivstellensatz

Under an additional assumption on the polynomials $g_1, \ldots, g_m$ defining the set $\mathbf{K}$ in (1.1), Putinar [84] shows an analogue of Schmüdgen's Positivstellensatz, which only involves $m+1$ sums of squares of polynomials instead of $2^m$ sums of squares of polynomials in Schmüdgen's Positivstellensatz.

The *quadratic module* generated by the polynomials $g_1, \ldots, g_m$ is defined as:

$$\mathcal{M}(g) := \left\{ \sigma_0 + \sum_{i=1}^m \sigma_i g_i \; : \; \sigma_i \in \Sigma[x], i = 0, 1, \ldots, m \right\}.$$

The quadratic module $\mathcal{M}(g)$ is called *Archimedean* if

$$\exists R > 0 \text{ s.t. } R^2 - \sum_{i=1}^n x_i^2 \in \mathcal{M}(g).$$

Note that the Archimedean assumption implies that $\mathbf{K}$ is compact, since it is contained in the ball $\mathbf{B}_R(0) := \{x \in \mathbb{R}^n : \|x\| \leq R\}$.

Then *Putinar's Positivstellensatz* can be stated as follows.

**Theorem 1.9.** *For the set* $\mathbf{K} = \{x \in \mathbb{R}^n : g_1(x) \geq 0, \ldots, g_m(x) \geq 0\}$, *assume that the quadratic module* $\mathcal{M}(g)$ *is Archimedean. For any polynomial* $f \in \mathbb{R}[x]$, *if* $f$ *is strictly positive on* $\mathbf{K}$, *then* $f$ *can be written as*

$$f = \sigma_0 + \sum_{i=1}^m \sigma_i g_i, \quad \text{where } \sigma_i \in \Sigma[x],\ i = 0, 1, \ldots, m.$$

Then, for any integer $r \geq 1$, the *truncated quadratic module* of degree $2r$, denoted as $\mathcal{M}_r(g)$, is defined as the subset of $\mathcal{M}(g)$ where the sums of squares of polynomials $\sigma_0, ..., \sigma_m$ meet the additional degree conditions:

$$\deg(\sigma_0) \leq 2r, \ \deg(\sigma_i g_i) \leq 2r \ (i = 1, \ldots, m).$$

Lasserre [47] introduces the following hierarchy of lower bounds for $f_{\min,\mathbf{K}}$

$$f_{las}^{(r)} := \sup\{\lambda : f - \lambda \in \mathcal{M}_r(g)\},$$

whose convergence to the global minimum $f_{\min,\mathbf{K}}$ (as $r$ increases) is guaranteed by Theorem 1.9.

One can easily see that $\mathcal{M}_r(g) \subseteq \mathcal{T}_{2r}(g)$, which implies $f_{las}^{(r)} \leq f_{sch}^{(2r)} \leq f_{\min,\mathbf{K}}$. However, the Schmüdgen bounds are more expensive to compute. Indeed, for each fixed $r$, one can compute the parameter $f_{las}^{(r)}$ via a semidefinite program, involving $m+1$ positive semidefinite matrices of order at most $\binom{n+r}{r}$, while computing the parameter $f_{sch}^{(2r)}$ needs solving a semidefinite program with $2^m$ positive semidefinite matrices of order at most $\binom{n+r}{r}$.

Lasserre's hierarchy has some nice properties. For instance, it exhibits finite convergence (i.e., $f_{las}^{(r)} = f_{\min,\mathbf{K}}$ holds for some $r$), for some classes of convex polynomial optimization problems (see Lasserre [51] and De Klerk and Laurent [19]). Moreover, finite convergence also holds when the description of $\mathbf{K}$ includes polynomial equations admitting only finitely many real solutions (see Laurent [57] and Nie [76]). Recently, Nie [75] shows that, under the Archimedean condition, Lasserre's hierarchy has finite convergence generically. Hence, finite convergence holds except for a set of data polynomials of Lebesgue measure zero. Nie and Schweighofer [74] show the following result about the quality of the bound $f_{las}^{(r)}$.

**Theorem 1.10.** *[74, Theorems 6 and 8] Assume the set* $\mathbf{K}$ *in* (1.1) *is contained in* $(-1,1)^n$ *and consider the parameter* $L(f)$ *from* (1.6)*. Then, there exist integers* $c, c' > 0$ *satisfying the following properties:*

*(i) Every polynomial* $f$ *of degree* $d$ *which is positive on* $\mathbf{K}$ *belongs to* $\mathcal{M}_r(g)$ *for some integer* $r$ *satisfying*

$$r \leq c \exp\left(\left(d^2 n^d \frac{L(f)}{f_{\min,\mathbf{K}}}\right)^c\right).$$

*(ii) For every polynomial* $f$ *of degree* $d$ *and for all integers* $r > c' \exp\left((2d^2n^d)^{c'}\right)$*, we have*

$$f - f_{\min,\mathbf{K}} + \frac{6d^3n^{2d}L(f)}{\sqrt[c']{\log\frac{r}{c'}}} \in \mathcal{M}_r(g), \ \textit{and thus } f_{\min,\mathbf{K}} - f_{las}^{(r)} \leq \frac{6d^3n^{2d}L(f)}{\sqrt[c']{\log\frac{r}{c'}}}.$$

For more information about Lasserre's hierarchy and its applications, see, e.g., [52, 55, 58, 31] and the references therein.

### 1.2.2 Optimization over measures

One can also reformulate polynomial optimization problems as optimization problems over measures, as introduced by Lasserre [47]. Assume $\mathbf{K}$ is compact. For computing the parameter $f_{\min,\mathbf{K}}$, the basic, fundamental idea of Lasserre [47] is to reformulate the problem as a minimization problem over the set $\mathcal{M}(\mathbf{K})$ of probability measures on the set $\mathbf{K}$. Namely,

$$f_{\min,\mathbf{K}} = \min_{\mu \in \mathcal{M}(\mathbf{K})} \mathbb{E}_\mu(f), \tag{1.14}$$

where

$$\mathbb{E}_\mu(f) := \int_{\mathbf{K}} f(x)\mu(dx) \tag{1.15}$$

denotes the expected value of $f$ over the probability measure $\mu$.

The above identity (1.14) is simple. As $f(x) \geq f_{\min,\mathbf{K}}$ for all $x \in \mathbf{K}$, one can integrate both sides with respect to any measure $\mu \in \mathcal{M}(\mathbf{K})$, which gives the inequality $\min_{\mu\in\mathcal{M}(\mathbf{K})} \int_{\mathbf{K}} f(x)\mu(dx) \geq f_{\min,\mathbf{K}}$. For the reverse inequality, let $\mu^*$ be the Dirac measure at a global minimizer $x^*$ of $f$ over $\mathbf{K}$, so that $\int_{\mathbf{K}} f(x)\mu^*(dx) = f(x^*) = f_{\min,\mathbf{K}} \geq \min_{\mu\in\mathcal{M}(\mathbf{K})} \int_{\mathbf{K}} f(x)\mu(dx)$.

Thus, in order to upper bound $f_{\min,\mathbf{K}}$ it suffices to choose a suitable probability measure on the set $\mathbf{K}$.

Later in this thesis we will investigate this approach which we will apply, in particular, to fixed-degree polynomial optimization over the standard simplex. We will consider some upper bounds, obtained by selecting some discrete probability distributions over the standard simplex. The multinomial distribution is used in Chapter 2 to give a much simplified convergence analysis for a known hierarchy of bounds, and the multivariate hypergeometric distribution is used in Chapter 3 to show a sharper rate of convergence.

Additionally, Lasserre [53] shows the following result, which roughly speaking says that, in (1.14), we may restrict to the Lebesgue measure with an arbitrary sum of squares of polynomials density function.

**Theorem 1.11.** *[53, Theorem 3.2] Let $\mathbf{K} \subseteq \mathbb{R}^n$ be compact and let $f$ be a continuous function on $\mathbb{R}^n$. Then the minimum of $f$ over $\mathbf{K}$ can be expressed as*

$$f_{\min,\mathbf{K}} = \inf_{h\in\Sigma[x]} \int_{\mathbf{K}} h(x)f(x)dx \quad s.t. \int_{\mathbf{K}} h(x)dx = 1.$$

By adding degree constraints on the polynomial density $h$ we get a hierarchy of upper bounds for $f_{\min,\mathbf{K}}$. That is, we obtain the upper bound

$$\underline{f}_{\mathbf{K}}^{(r)} := \inf_{h\in\Sigma[x]_r} \int_{\mathbf{K}} h(x)f(x)dx \quad \text{s.t.} \int_{\mathbf{K}} h(x)dx = 1, \tag{1.16}$$

where $\Sigma[x]_r$ denotes the set of sums of squares of polynomials with degree at most $2r$.

Obviously, one has

$$f_{\min,\mathbf{K}} \leq \cdots \leq \underline{f}_{\mathbf{K}}^{(r+1)} \leq \underline{f}_{\mathbf{K}}^{(r)} \leq \cdots \leq \underline{f}_{\mathbf{K}}^{(1)},$$

and $\lim_{r\to\infty} \underline{f}_{\mathbf{K}}^{(r)} = f_{\min,\mathbf{K}}$ holds by Theorem 1.11.

Moreover, if we know the explicit values of the moments $\int_{\mathbf{K}} x^\alpha dx$ for any $\alpha \in \mathbb{N}^n$ (which holds, e.g., when $\mathbf{K}$ is a full-dimensional simplex, hypercube, or a Euclidean ball), then we can compute $\underline{f}_{\mathbf{K}}^{(r)}$ by solving a semidefinite program.

In Chapter 5 we will analyze the quality of this hierarchy of upper bounds, and show that its rate of convergence satisfies $\underline{f}_{\mathbf{K}}^{(r)} - f_{\min,\mathbf{K}} = O(\frac{1}{\sqrt{r}})$.

## 1.3 Notation

In this section we collect all notation we use in this thesis.

### 1.3.1 Sets

We use $\mathbb{R}$, $\mathbb{R}_+$, $\mathbb{Q}$, $\mathbb{Z}$ and $\mathbb{N}$ to denote the sets of real numbers, nonnegative real numbers, rational numbers, integers, and nonnegative integers, respectively, and we use $\mathbb{R}^n$, $\mathbb{R}^n_+$, $\mathbb{Q}^n$, $\mathbb{Z}^n$ and $\mathbb{N}^n$ to denote the corresponding sets of $n$-dimensional vectors.

Given a finite set $V$ and an integer $t$, $\mathcal{P}(V)$ denotes the collection of all subsets of $V$, $\mathcal{P}_t(V) := \{I \subseteq V : |I| \leq t\}$, and $\mathcal{P}_{=t}(V) := \{I \subseteq V : |I| = t\}$. We denote $[n] = \{1, 2, \ldots, n\}$.

For two vectors $\alpha, \beta \in \mathbb{N}^n$, the inequality $\alpha \leq \beta$ is coordinate-wise and means that $\alpha_i \leq \beta_i$ for any $i \in [n]$. The support of $x \in \mathbb{R}^n$ is the set $\{i \in [n] : x_i \neq 0\}$. For $x \in \mathbb{R}^n$ and $S \subseteq [n]$, we denote $x(S) := \sum_{i\in S} x_i$. We let $\mathbf{e}$ denote the all-ones vector and let $\mathbf{e}_1, \ldots, \mathbf{e}_n$ denote the standard unit vectors. For $I \subseteq [n]$ we set $\mathbf{e}_I := \sum_{i\in I} \mathbf{e}_i$, and use $|I|$ to denote the cardinality of $I$.

Throughout, we let

$$\mathbf{Q}_n = [0, 1]^n$$

denote the $n$-dimensional unit hypercube and

$$\mathbf{B}_\epsilon(a) = \{x \in \mathbb{R}^n : \|x - a\| \leq \epsilon\}$$

denote the Euclidean ball centered at $a \in \mathbb{R}^n$ with radius $\epsilon > 0$. Moreover, the sets

$$\Delta_n = \{x \in \mathbb{R}^n_+ : \sum_{i=1}^n x_i = 1\}$$

and

$$\widehat{\Delta_n} := \{x \in \mathbb{R}^n_+ : \sum_{i=1}^n x_i \leq 1\}$$

denote, respectively, the standard simplex and the full-dimensional simplex in $\mathbb{R}^n$. Given an integer $r \geq 1$, define

$$I(n, r) = \{x \in \mathbb{N}^n : \sum_{i=1}^n x_i = r\},$$

$$\Delta(n, r) = \{x \in \Delta_n : rx \in \mathbb{N}^n\},$$

and

$$N(n, r) = \{x \in \mathbb{N}^n : \sum_{i=1}^n x_i \leq r\}.$$

The set of symmetric $n \times n$ matrices is denoted as $\mathcal{S}_n$. A matrix $A \in \mathcal{S}_n$ is positive semidefinite (resp., copositive) if $x^T Ax \geq 0$ for all $x \in \mathbb{R}^n$ (resp., $x^T Ax \geq 0$ for all $x \geq 0$). Then, $\mathcal{S}_n^+$ denotes the set of $n \times n$ positive semidefinite matrices, and $\mathcal{C}_n$ is the set of $n \times n$ copositive matrices.

### 1.3.2 Polynomials and functions

Let $\mathbb{R}[x] = \mathbb{R}[x_1, \ldots, x_n]$ denote the set of multivariate polynomials in $n$ variables with real coefficients. We denote monomials in $\mathbb{R}[x]$ as $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ for $\alpha \in \mathbb{N}^n$, with degree $|\alpha| = \sum_{i=1}^n \alpha_i$. For a polynomial $f = \sum_{\alpha \in \mathbb{N}^n} f_\alpha x^\alpha$, its degree is defined as $\deg(f) = \max_{\{\alpha : f_\alpha \neq 0\}} |\alpha|$, and $f$ is called homogeneous if all its monomials have the same degree. Furthermore, we set $\phi_\alpha(x) := x^\alpha$.

Let $\mathbb{R}_+[x]$ denote the set of polynomials with nonnegative real coefficients. For an integer $r \geq 1$, $\mathbb{R}[x]_r$ denotes the set of polynomials of degree at most $r$, and $\mathbb{R}_+[x]_r$ consists of all polynomials with nonnegative real coefficients of degree at most $r$.

$\Sigma[x]$ is the set of sums of squares of polynomials, and $\Sigma[x]_r$ consists of all sums of squares of polynomials with degree at most $2r$. Moreover, let $\mathcal{H}_{n,d}$ denote the set of all multivariate real homogeneous polynomials in $n$ variables with degree $d$.

The monomial $x^\alpha$ is square-free (or multilinear) if $\alpha \in \{0,1\}^n$ and a polynomial $f$ is square-free if all its monomials are square-free. For $I \subseteq [n]$, we use the notation $x^I := \prod_{i\in I} x_i$. Hence, a square-free polynomial $f$ can be written as $f = \sum_{I\subseteq[n]} f_I x^I$.

Given a set $\mathbf{K} \subseteq \mathbb{R}^n$, we say that $f$ is positive (resp., nonnegative) on $\mathbf{K}$ when $f(x) > 0$ (resp., $f(x) \geq 0$) for all $x \in \mathbf{K}$. Furthermore, we denote $\mathscr{P}(\mathbf{K})$ as the set of polynomials that are nonnegative on $\mathbf{K}$. Given a set $\mathbf{K} \subseteq \mathbb{R}^n$, we use $w_{\min}(\mathbf{K})$ to denote the minimal width of $\mathbf{K}$, which is defined as the minimum distance between two distinct parallel supporting hyperplanes of $\mathbf{K}$, and we use $D(\mathbf{K}) = \sup_{x,y\in\mathbf{K}} \|x-y\|^2$ to denote the (squared) diameter of $\mathbf{K}$, where $\|x\| = \sqrt{\sum_{i=1}^n x_i{}^2}$ is the $\ell_2$-norm.

For $x \in \mathbb{R}$ and $d \in \mathbb{N}$, we denote $x^{\underline{d}} = x(x-1)(x-2)\cdots(x-d+1)$ and thus $x^{\underline{d}} = 0$ if $x$ is an integer with $0 \leq x \leq d-1$. For $x \in \mathbb{R}^n$ and $\alpha \in \mathbb{N}^n$, we denote $x^{\underline{\alpha}} = \prod_{i=1}^n x_i^{\underline{\alpha_i}}$. For $\alpha \in \mathbb{N}^n$, we denote $\alpha! = \alpha_1!\alpha_2!\cdots\alpha_n!$.

We use $\Gamma(\cdot)$ to denote the Euler gamma function. For integers $n, k \in \mathbb{N}$, the Stirling number of the second kind $S(n,k)$ counts the number of ways of partitioning a set of $n$ objects into $k$ nonempty subsets. Thus $S(n,k) = 0$ if $k > n$, $S(n,0) = 0$ if $n \geq 1$, and $S(0,0) = 1$ by convention. For any integer $k \geq -1$, the double factorial $k!!$ is defined as

$$k!! = \begin{cases} k\cdot(k-2)\cdots 3\cdot 1, & \text{if } k > 0 \text{ is odd},\\ k\cdot(k-2)\cdots 4\cdot 2, & \text{if } k > 0 \text{ is even},\\ 1 & \text{if } k = 0 \text{ or } k = -1. \end{cases}$$

Let $f(x)$, $g(x)$: $\mathbb{R} \to \mathbb{R}$ be two non-negative real-valued functions. We write $f(x) = O(g(x))$ if there exist positive numbers $M$ and $x_0$ such that $f(x) \leq Mg(x)$ for all $x \geq x_0$. Moreover, we write $f(x) = \Omega(g(x))$ if there exist positive numbers $M$ and $x_0$ such that $f(x) \geq Mg(x)$ for all $x \geq x_0$; see, e.g., [69, Definition B.1].

### 1.3.3 Graphs

Given a graph $G = (V,E)$, $\overline{G} = (V, \overline{E})$ denotes its complementary graph whose edges are the pairs of distinct nodes $i, j \in V(G)$ with $ij \notin E$. Throughout we also set $V = V(G)$, $E = E(G)$ and we always assume $V(G) = [n]$. $K_n$ denotes the complete graph on $n$ nodes, and $C_n$ denotes the circuit on $n$ nodes.

A set $S \subseteq V$ is stable (or independent) if no two distinct nodes of $S$ are adjacent in $G$ and a clique in $G$ is a set of pairwise adjacent nodes. The maximum cardinality of a stable set (resp., clique) in $G$ is denoted by $\alpha(G)$ (resp., $\omega(G)$); thus $\omega(G) = \alpha(\overline{G})$. The *chromatic number* $\chi(G)$ is the minimum number of colors needed to color the nodes of $G$ in such a way that adjacent nodes receive distinct colors.

For a node $i \in V$, $G - i$ denotes the graph obtained by deleting node $i$ from $G$, and $G \ominus i$ denotes the graph obtained from $G$ by removing $i$ as well as the set $N(i)$ of its neighbours. For $U \subseteq V$, $G \backslash U$ denotes the graph obtained by deleting all nodes of $U$. For an edge $e \in E$, let $G \backslash e$ denote the graph obtained by deleting edge $e$ from $G$, and let $G/e$ denote the graph obtained from $G$ by contracting edge $e$. Consider two graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ such that $V_1 \cap V_2$ is a clique of cardinality $t$ in both $G_1$ and $G_2$. Then the graph $G = (V_1 \cup V_2, E_1 \cup E_2)$ is called the *clique t-sum* of $G_1$ and $G_2$.

## 1.4 Contents of the thesis

The rest of this thesis is divided into three parts. In what follows, I elaborate about the contents of this thesis in the three parts.

### 1.4.1 Polynomial optimization over the standard simplex

In Part I, we consider the problem of minimizing a polynomial over the standard simplex, i.e., the problem of computing $f_{\min,\Delta_n}$. A well studied approach to approximate $f_{\min,\Delta_n}$ is to consider the hierarchy of upper bounds, obtained by minimizing over the set of regular grid points in the standard simplex, with a given denominator. That is, consider the parameters $f_{\Delta(n,r)}$ as defined in (1.9).

For any homogeneous polynomial $f \in \mathcal{H}_{n,d}$, De Klerk et al. [21] study the parameter $f_{\Delta(n,r)}$ and show that its *convergence ratio*

$$\rho_r(f) := \frac{f_{\Delta(n,r)} - f_{\min,\Delta_n}}{f_{\max,\Delta_n} - f_{\min,\Delta_n}} \tag{1.17}$$

satisfies

$$\rho_r(f) \leq \frac{C(d)}{r}, \tag{1.18}$$

where $C(d)$ is a constant depending only on $d$ (see Theorem 2.1 for details). Observe that the parameter $f_{\Delta(n,r)}$ can be calculated via $|\Delta(n,r)| = \binom{n+r-1}{r}$ evaluations of $f$. Hence, it can be computed in polynomial time for any fixed $r$. Thus the parameters

$f_{\Delta(n,r)}$ with increasing $r$ lead to a polynomial time approximation scheme (PTAS, see Definition 2.2) for fixed-degree polynomial optimization.

In Chapter 2, we give a much simplified proof for the inequality in (1.18). The idea for our new proof can be described as follows. As in (1.14), we can reformulate $f_{\Delta(n,r)}$ as optimization over measures. That is,

$$f_{\Delta(n,r)} = \min_{\mu \in \mathcal{M}(\Delta(n,r))} \mathbb{E}_\mu(f),$$

where $\mathbb{E}_\mu(f)$ is defined in (1.15).

Then our strategy is to study an upper bound for $f_{\Delta(n,r)}$, obtained by choosing the multinomial distribution as the probability measure on $\Delta(n,r)$. It turns out that this upper bound is closely related to Bernstein approximation, which is a classical tool in approximation theory. Namely, the upper bound boils down to the Bernstein approximation of $f$ over the standard simplex. Then the convergence analysis is based on using some properties of Bernstein approximation. Moreover, our analysis completes the analysis of the random walk approach proposed by Nesterov [72] to upper bound the parameter $f_{\Delta(n,r)}$.

Then, we show in Chapter 3 that by using another distribution on $\Delta(n,r)$, the multivariate hypergeometric distribution, we can sharpen the analysis for the convergence of $f_{\Delta(n,r)}$. To be more precise, we show that under some conditions on $f$,

$$\rho_r(f) \leq \frac{C(f)}{r^2}, \tag{1.19}$$

where the constant $C(f)$ depends on the polynomial $f$ but not on $r$. Namely, this result holds for the quadratic case (i.e., when $f$ is quadratic), and it also holds in the general case assuming the existence of a rational global minimizer. However, the best-known upper estimates for $C(f)$ are exponential in $n$ in general, which means that the estimate in (1.19) does not yield a PTAS for the problem of minimizing a quadratic polynomial over the standard simplex.

In addition, in Chapter 4 we consider the upper bound $f_{\Delta(n,r)}$ together with the lower bound $f_{\min}^{(r-d)}$, which we introduced earlier in (1.8). We uncover their mutual relationship and give refined upper bounds for the range $f_{\Delta(n,r)} - f_{\min}^{(r-d)}$ in terms of the range $f_{\max,\Delta_n} - f_{\min,\Delta_n}$.

### 1.4.2 Polynomial optimization over a compact set

In Part II we investigate the more general problem of minimizing a continuous function over a compact set. We focus on the hierarchy of upper bounds $\underline{f}_{\mathbf{K}}^{(r)}$ for $f_{\min,\mathbf{K}}$, defined as in (1.16):

$$\underline{f}_{\mathbf{K}}^{(r)} = \inf_{h \in \Sigma[x]_r} \int_{\mathbf{K}} h(x) f(x) dx \quad \text{s.t.} \int_{\mathbf{K}} h(x) dx = 1.$$

When $f$ is a polynomial, this hierarchy has been investigated in [47, 53]. In particular, for fixed $r$, the parameter $\underline{f}_{\mathbf{K}}^{(r)}$ can be computed in polynomial time for some cases, e.g., when $\mathbf{K}$ is a full-dimensional simplex, hypercube, or a Euclidean ball. However, no information about its convergence rate is known.

In Chapter 5, we show that its convergence rate is in $O(1/\sqrt{r})$. More precisely, we prove that

$$\underline{f}_{\mathbf{K}}^{(r)} - f_{\min,\mathbf{K}} \le \frac{\zeta(\mathbf{K}) M_f}{\sqrt{r}} \quad \text{for any } r \text{ large enough},$$

where $\zeta(\mathbf{K})$ is a constant depending only on $\mathbf{K}$, and $M_f$ is the Lipschitz constant of $f$ on $\mathbf{K}$ (see Theorem 5.7 for details). Our result applies to the case when $f$ is Lipschitz continuous and $\mathbf{K}$ is a full-dimensional compact set satisfying some geometrical condition (which is satisfied, e.g., for any full-dimensional compact convex set). The main idea is to use the Taylor series of the Gaussian distribution function truncated at degree $2r$ as the sum of squares density function in order to carry out the analysis.

In addition, we indicate how to sample feasible points in $\mathbf{K}$ from the probability distribution defined by the optimal density function $h^*$, obtained as the optimal solution of the program (1.16). We also present numerical results for several polynomial test functions on the hypercube. In these examples, we observe that the sampling based on $h^*$ generates 'better' feasible solutions than the uniform sampling from $\mathbf{K}$.

### 1.4.3 An application in graph theory

In part III we consider the maximum stable set problem in graph theory. In particular, we analyze the following formulation for $\alpha(G)$ considered by Park and Hong [78]: given a graph $G = (V, E)$, its stability number $\alpha(G)$ can be computed via the following quadratic maximization problem on the hypercube:

$$\alpha(G) = \max_{x \in [0,1]^{|V|}} \sum_{i \in V} x_i - \sum_{ij \in E} x_i x_j. \tag{1.20}$$

Hence, we can use the representation result of Handelman [38], as explained earlier in Section 1.2.1, to build a hierarchy of upper bounds for $\alpha(G)$. It turns out that this hierarchy converges in finitely many steps. Then we focus on the smallest number of steps needed for the finite convergence, which is called the Handelman rank (see Definition 6.10). More precisely, we consider the following question: given a graph, what is its Handelman rank?

We relate the Handelman rank with structural properties of graphs. In particular, we use fractional clique covers to upper bound the Handelman rank for perfect graphs and we determine its exact value in the vertex-transitive case. Moreover, we show two upper bounds on the Handelman rank in terms of the (fractional) stability number of the graph, and we compute the Handelman rank for several classes of graphs including odd circuits and wheels and their complements. We also point out links to several other classical hierarchies of bounds by Sherali-Adams, Lovász-Schrijver, Lasserre and De Klerk-Pasechnik. Additionally, we give an explicit formulation for the Handelman hierarchy applied to the max-cut problem in terms of valid inequalities of the cut polytope.

# Publications and preprints

The rest of this thesis includes five chapters, which are based on the following publications and preprints:

| | |
|---|---|
| Chapter 2 | [22] An alternative proof of a PTAS for fixed-degree polynomial optimization over the simplex. de Klerk, E., Laurent, M., Sun, Z. *Math. Program. (online first)*, DOI: 10.1007/s10107-014-0825-6 (2014). |
| Chapter 3 | [23] An error analysis for polynomial optimization over the simplex based on the multivariate hypergeometric distribution. de Klerk, E., Laurent, M., Sun, Z. (2014) *SIAM J. Optim.* (accepted with minor revision) |
| Chapter 4 | [92] A refined error analysis for fixed-degree polynomial optimization over the simplex. Sun, Z. *J. Oper. Res. Soc. China*, 2(3) pp 379–393 (2014). |
| Chapter 5 | [24] Convergence analysis for Lasserre's measure-based hierarchy of upper bounds for polynomial optimization. de Klerk, E., Laurent, M., Sun, Z. (2014) Preprint at arXiv: 1411.6867 |
| Chapter 6 | [59] Handelman's hierarchy for the maximum stable set problem. Laurent, M., Sun, Z. *J. Global Optim.*, 60(3) pp 393–423 (2014). |

# Part I

# Polynomial Optimization over the Standard Simplex

In this Part, we consider the problem of minimizing a polynomial $f \in \mathbb{R}[x]$ on the standard simplex

$$\Delta_n = \left\{ x \in \mathbb{R}^n_+ : \sum_{i=1}^n x_i = 1 \right\}.$$

That is, the problem of computing the parameter

$$f_{\min,\Delta_n} = \min_{x \in \Delta_n} f(x). \tag{1.21}$$

As we have mentioned before, this problem is NP-hard, even if $f$ is a quadratic function, as it contains the maximum stable set problem (1.2) as a special case. For more information about the complexity of optimization over the simplex, see, e.g., [16, 17].

Observe that one can assume w.l.o.g. that $f$ is homogeneous (say, of degree $d$). Indeed, if $f = \sum_{s=0}^d f_s$, where $f_s$ is homogeneous of degree $s$, then $\min_{x\in\Delta_n} f(x) = \min_{x\in\Delta_n} \tilde{f}(x)$ after setting $\tilde{f} = \sum_{s=0}^d f_s \left(\sum_{i=1}^n x_i\right)^{d-s}$.

We focus on the bound

$$f_{\Delta(n,r)} = \min \ f(x) \text{ s.t. } x \in \Delta(n,r) = \{x \in \Delta_n : rx \in \mathbb{N}^n\},$$

which was defined in (1.9).

Error bounds for $f_{\Delta(n,r)}$ have been shown by De Klerk and Bomze [8] (for quadratic polynomial $f$), and by De Klerk et al. [21] (for general polynomial $f$). They show that the convergence ratio of $f$

$$\rho_r(f) = \frac{f_{\Delta(n,r)} - f_{\min,\Delta_n}}{f_{\max,\Delta_n} - f_{\min,\Delta_n}}$$

(as defined in (1.17)) satisfies

$$\rho_r(f) \le \frac{C(d)}{r},$$

where $C(d)$ is a constant depending only on $d$, see Theorem 2.1 for details.

In Chapter 2, we give a new proof for the above inequality, and we also refine the known constant $C(d)$ in the case $d = 3$. For the proof, we first reformulate $f_{\Delta(n,r)}$ as

$$f_{\Delta(n,r)} = \min_{\mu \in \mathcal{M}(\Delta(n,r))} \mathbb{E}_\mu(f),$$

where $\mathcal{M}(\Delta(n,r))$ denotes the set of probability measures on $\Delta(n,r)$ and $\mathbb{E}_\mu(f) = \int_{\Delta(n,r)} f(x)\mu(dx)$ denotes the expected value of $f$ over the probability measure $\mu$.

Then the main idea is to study an upper bound for $f_{\Delta(n,r)}$, obtained by choosing the multinomial distribution as the probability measure on $\Delta(n,r)$. It turns out that this upper bound is equal to the Bernstein approximation of $f$ over the standard simplex, and the convergence analysis uses some properties of Bernstein approximation. Moreover, our analysis in Chapter 2 is closely related to Nesterov's random walk on $\Delta(n,r)$ in [72]. However, Nesterov [72] considers only polynomials of degree at most 3 and square-free polynomials. Hence, we complete his analysis for general polynomials by placing it in the well-studied framework of Bernstein approximation and clarifying the link to the multinomial distribution.

In Chapter 2 several examples are investigated and it turns out that $\rho_r(f) = O(1/r^2)$ holds for all of them, which is sharper than the $O(1/r)$ proved bound. Thus an open question raises: does $\rho_r(f) = O(1/r^2)$ hold in general?

In Chapter 3 we show that by using another distribution on $\Delta(n,r)$, the multivariate hypergeometric distribution, we can show that under some conditions on the polynomial $f$,

$$\rho_r(f) \leq \frac{C(f)}{r^2} \tag{1.22}$$

holds, where $C(f)$ depends on the polynomial $f$. More precisely, this result holds for the quadratic case (i.e., when $f$ is quadratic), and also holds in the general case assuming the existence of a rational global minimizer. However, the best-known upper bounds on $C(f)$ are exponential in $n$ in general.

Finally, in Chapter 4 we consider $f_{\Delta(n,r)}$, together with the parameter

$$f_{\min}^{(r-d)} = \sup \lambda \ \text{ s.t. } \ \left(\sum_{i=1}^n x_i\right)^{r-d} \left(f - \lambda \left(\sum_{i=1}^n x_i\right)^d\right) \in \mathbb{R}_+[x],$$

defined as in (1.8), which is a lower bound for $f_{\min,\Delta_n}$ obtained from Pólya's theorem (Theorem 1.1). In fact, both $f_{\Delta(n,r)}$ and $f_{\min}^{(r-d)}$ have been studied in the literature. In particular, De Klerk et al. [21] show upper bounds for $f_{\Delta(n,r)} - f_{\min,\Delta_n}$ and $f_{\min,\Delta_n} - f_{\min}^{(r-d)}$ separately. We show upper bounds for $f_{\Delta(n,r)} - f_{\min}^{(r-d)}$ and refine the previous known upper bounds, obtained by adding up the upper bounds for $f_{\Delta(n,r)} - f_{\min,\Delta_n}$ and $f_{\min,\Delta_n} - f_{\min}^{(r-d)}$.

# Chapter 2

# New proof for a polynomial time approximation scheme (PTAS)

## 2.1 Introduction

For the problem of computing $f_{\min,\Delta_n}$, many approximation methods have been studied in the literature. In fact, when $f$ has fixed degree $d$, there is a polynomial time approximation scheme (PTAS, see Definition 2.2 below) for this problem, as is shown by Bomze and De Klerk [8] (for quadratic $f$), and by De Klerk, Laurent and Parrilo [21] (for general fixed-degree $f$). The PTAS is particularly simple: it takes the minimum of $f$ on the regular grid

$$\Delta(n,r) = \{x \in \Delta_n : rx \in \mathbb{N}^n\}$$

for increasing values of $r$. Recall that we denote the minimum over the grid by

$$f_{\Delta(n,r)} = \min_{x \in \Delta(n,r)} f(x).$$

Hence, the computation of $f_{\Delta(n,r)}$ requires $|\Delta(n,r)| = \binom{n+r-1}{r}$ evaluations of $f$, which is polynomial in $n$ for fixed $r$.

Several properties of the regular grid $\Delta(n,r)$ have been studied in the literature. In Bos [10], the Lebesgue constant of $\Delta(n,r)$ is studied in the context of Lagrange interpolation and finite element methods. Given a point $x \in \Delta_n$, Bomze, Gollowitzer and Yildirim [9] study a scheme to find the closest point to $x$ on $\Delta(n,r)$ with respect to certain norms (including $\ell_p$-norms for finite $p$). Furthermore, as the sequence of $f_{\Delta(n,r)}$ may not be monotone non-increasing for increasing values of $r$, Sagol and

Yildirim [86] and Yildirim [100] consider the parameter $\min_{x \in \cup_{k=2}^{r} \Delta(n,k)} f(x)$ (which is monotone non-increasing for increasing values of $r$) for homogeneous quadratic polynomial, and analyze its quality.

The following error bounds are known for the approximation $f_{\Delta(n,r)}$ of $f_{\min,\Delta_n}$.

**Theorem 2.1** ((i) Bomze-De Klerk [8] and (ii) De Klerk-Laurent-Parrilo [21])**.**

*(i) For any quadratic polynomial* $f \in \mathcal{H}_{n,2}$ *and* $r \geq 2$*, one has*

$$f_{\Delta(n,r)} - f_{\min,\Delta_n} \leq \frac{f_{\max,\Delta_n} - f_{\min,\Delta_n}}{r}.$$

*(ii) For any polynomial* $f \in \mathcal{H}_{n,d}$ *and* $r \geq d$*, one has*

$$f_{\Delta(n,r)} - f_{\min,\Delta_n} \leq \left(1 - \frac{r^{\underline{d}}}{r^d}\right) \binom{2d-1}{d} d^d (f_{\max,\Delta_n} - f_{\min,\Delta_n}),$$

*where* $r^{\underline{d}} = r(r-1)\cdots(r-d+1)$.

Note that $1 - \frac{r^{\underline{d}}}{r^d} = O(\frac{1}{r})$, and thus the above results imply the existence of a PTAS in the sense of the following definition, that has been used by several authors (see, e.g., [5, 17, 21, 73, 96]).

**Definition 2.2.** *[PTAS] Given any compact set* $\mathbf{K}$*, a value* $\psi_\epsilon$ *approximates* $f_{\min,\mathbf{K}}$ *with relative accuracy* $\epsilon$ *in* $[0,1]$ *if*

$$|\psi_\epsilon - f_{\min,\mathbf{K}}| \leq \epsilon(f_{\max,\mathbf{K}} - f_{\min,\mathbf{K}}).$$

*The approximation is called implementable if* $\psi_\epsilon = f(x_\epsilon)$ *for some feasible* $x_\epsilon$*. If a problem allows an implementable approximation* $\psi_\epsilon = f(x_\epsilon)$ *for each* $\epsilon \in (0,1]$*, such that the feasible* $x_\epsilon$ *can be computed in time polynomial in* $n$ *and the bit size required to represent* $f$*, then we say that the problem allows a polynomial time approximation scheme (PTAS).*

The main contribution of this chapter is to provide new insight into the PTAS by establishing precise connections with the multinomial distribution and Bernstein approximation. More precisely, we give simplified proofs of the PTAS result. In particular, our proof for the quadratic case is completely elementary, and much simpler than the proof given in [8]. We also refine the relevant error bound in the special case of degree three polynomials. In addition, our analysis in this chapter is closely related to a probabilistic proof given by Nesterov [72], see Section 2.2.2 for precise connections.

## 2.2 Preliminaries

To analyze the quality of the parameter $f_{\Delta(n,r)}$, we start by reformulating $f_{\Delta(n,r)}$ as a minimization problem over the set of probability measures (as we saw earlier in (1.14)), i.e.,

$$f_{\Delta(n,r)} = \min_{\mu \in \mathcal{M}(\Delta(n,r))} \mathbb{E}_\mu(f), \tag{2.1}$$

where $\mathbb{E}_\mu(f) = \int_{\Delta(n,r)} f(x)\mu(dx)$.

Then we can obtain an upper bound for $f_{\Delta(n,r)}$ by setting the measure $\mu$ to be a suitable probability measure on the regular grid $\Delta(n,r)$. In this chapter we focus on the upper bound obtained by selecting the multinomial distribution with appropriate parameters as measure $\mu$. It turns out that this upper bound boils down to the Bernstein approximation of $f$ over the standard simplex $\Delta_n$. Moreover, our approach is closely related to Nesterov's random walk in the standard simplex [72].

Next we review some necessary background material on the multinomial distribution, Nesterov's random walk, and Bernstein approximation.

### 2.2.1 The multinomial distribution

Recall that the *multinomial distribution* with parameters $r, n$, and $x_1, \ldots, x_n$ (where $x \in \Delta_n$) can be explained by rolling a loaded dice. More precisely, consider a loaded dice with $n$ sides. We roll the dice $r$ times, and at each trial the probability of seeing $i$ is $x_i$. We let the random variable $Y_i$ denote the number of times that $i$ is seen.

Then, $Y = (Y_1, \ldots, Y_n)$ has the multinomial distribution, with parameters $r, n$, and $x_1, \ldots, x_n$ (where $x \in \Delta_n$). Given $\alpha \in I(n,r) = \{\alpha \in \mathbb{N}^n : \sum_{i=1}^n \alpha_i = r\}$, the probability of obtaining the outcome $Y = \alpha$, is equal to

$$\mathbf{Pr}\left[Y_1 = \alpha_1, \ldots, Y_n = \alpha_n\right] = \frac{r!}{\alpha!} x^\alpha, \qquad \alpha \in I(n,r). \tag{2.2}$$

Then the normalized random variable $X = \frac{1}{r}Y$ takes its values in $\Delta(n,r)$, and the expected value of $f(X)$ is

$$\mathbb{E}[f(X)] = \sum_{\alpha \in I(n,r)} f\left(\frac{\alpha}{r}\right) \frac{r!}{\alpha!} x^\alpha. \tag{2.3}$$

Since the random variable $X$ takes its values in $\Delta(n,r)$, this implies directly that the expected value of $f(X)$ is at least the minimum of $f$ over $\Delta(n,r)$. That is,

$$f_{\Delta(n,r)} \leq \mathbb{E}[f(X)]. \tag{2.4}$$

As we will see in (2.7) below, it turns out that $\mathbb{E}[f(X)]$ is equal to $B_r(f)(x)$, the Bernstein approximation of $f$ of order $r$ at the point $x \in \Delta_n$. Our new proof will be based on exploiting the properties of Bernstein approximation on the standard simplex.

On the other hand, as mentioned before, this analysis is closely related to Nesterov's random walk in the standard simplex proposed in [72]. Next we illustrate the precise connection.

### 2.2.2 Nesterov's random walk in the standard simplex

Nesterov [72] proposes an alternative probabilistic argument for estimating the quality of the bounds $f_{\Delta(n,r)}$. He considers a random walk on the standard simplex $\Delta_n$, which generates a sequence of random points $x^{(r)} \in \Delta(n,r)$ $(r = 1, 2, \dots)$. Thus the expected value $\mathbb{E}[f(x^{(r)})]$ of the evaluation of the polynomial $f$ at $x^{(r)}$ satisfies:

$$f_{\Delta(n,r)} \leq \mathbb{E}[f(x^{(r)})].$$

For completeness, we describe Nesterov's approach as follows.

Let $x \in \Delta_n$ and let $\zeta$ be a discrete random variable taking values in $\{1, \dots, n\}$ where the probability of the event $\zeta = i$ is given by $x_i$. That is,

$$\mathbf{Pr}[\zeta = i] = x_i \ (i = 1, \dots, n). \tag{2.5}$$

Consider the random process:

$$y^{(0)} = 0 \in \mathbb{R}^n, \ y^{(r)} = y^{(r-1)} + \mathbf{e}_{\zeta_r} \ (r \geq 1),$$

where $\zeta_r$ are independent random variables distributed according to (2.5). In other words, $y^{(r)}$ equals $y^{(r-1)} + \mathbf{e}_i$ with probability $x_i$. One can easily check that $y^{(r)}$ has the multinomial distribution, with parameters $r, n$ and $x_1, \dots, x_n$ (where $x \in \Delta_n$). Hence, by (2.2), for any given $\alpha \in I(n,r)$, the probability of the event $y^{(r)} = \alpha$ is given by

$$\mathbf{Pr}[y^{(r)} = \alpha] = \frac{r!}{\alpha!} x^\alpha.$$

Finally, define

$$x^{(r)} = \frac{1}{r} y^{(r)} \in \Delta(n,r) \quad (r \geq 1).$$

Thus one has

$$\mathbf{Pr}[x^{(r)} = \alpha/r] = \mathbf{Pr}[y^{(r)} = \alpha] = \frac{r!}{\alpha!} x^\alpha,$$

and it immediately follows that

$$\begin{aligned}\mathbb{E}[f(x^{(r)}] &= \sum_{\alpha\in I(n,r)} \mathbf{Pr}[x^{(r)} = \alpha/r]f(\alpha/r)\\ &= \sum_{\alpha\in I(n,r)} \frac{r!}{\alpha!}x^\alpha f\left(\frac{\alpha}{r}\right).\end{aligned} \tag{2.6}$$

Note that the value of $\mathbb{E}[f(x^{(r)})]$ in (2.6) is equal to the value of $\mathbb{E}[f(X)]$ in (2.3). Thus, in this sense, our approach using Bernstein approximation is equivalent to the random walk approach of Nesterov [72].

On the other hand, in [72] the link with Bernstein approximation is not made, and the author calculates the values $\mathbb{E}[f(x^{(r)})]$ from first principles for polynomials up to degree four and for square-free polynomials. Based on this Nesterov [72] gives the error bounds in Theorems 2.8 and 2.14 below for the quadratic and square-free cases. However, he does not consider the general case. Thus the analysis in this chapter completes the analysis in [72].

### 2.2.3 Bernstein approximation on the standard simplex

We now review some necessary background material for Bernstein approximation.

The Bernstein approximation of order $r \geq 1$ on the standard simplex of a continuous function $f$ is the polynomial $B_r(f) \in \mathcal{H}_{n,r}$ defined by

$$B_r(f)(x) = \sum_{\alpha\in I(n,r)} f\left(\frac{\alpha}{r}\right)\frac{r!}{\alpha!}x^\alpha, \tag{2.7}$$

where $\alpha! = \prod_{i=1}^n \alpha_i!$ and $x^\alpha = \prod_{i=1}^n x_i^{\alpha_i}$. For instance, for the constant polynomial $f \equiv 1$, its Bernstein approximation of any order $r$ is $\sum_{\alpha\in I(n,r)} \frac{r!}{\alpha!}x^\alpha$, which is equal to $(\sum_{i=1}^n x_i)^r$ by the multinomial theorem, and thus to 1 for any $x \in \Delta_n$.

There is a vast literature on Bernstein approximation, and the interested reader may consult, e.g., the papers by Ditzian [28, 29], Ditzian and Zhou [30], the book by Altomare and Campiti [2], and the references therein for more details than given here.

Here we state one well-known result that shows uniform convergence of the Bernstein approximations $B_r(f)$ to $f$ as $r$ increases to infinity.

**Theorem 2.3** (See, e.g. [2], Section 5.2.11)**.** *Let $f : \mathbb{R}^n \to \mathbb{R}$ be any continuous function defined on $\Delta_n$ and let $B_r(f)$ be as defined in (2.7). One has*

$$|B_r(f)(x) - f(x)| \leq 2\omega\left(f, \frac{1}{\sqrt{r}}\right) \qquad \forall x \in \Delta_n,$$

*where $\omega$ denotes the modulus of continuity:*

$$\omega(f,\delta) := \max_{\substack{x,y\in\Delta_n \\ \|x-y\|\leq\delta}} |f(x) - f(y)| \qquad (\delta \geq 0).$$

Next we state some simple inequalities relating a polynomial, its Bernstein approximation and their minimum over the set $\Delta(n,r)$ of grid points.

**Lemma 2.4.** *Given a polynomial $f \in \mathcal{H}_{n,d}$ and $r \geq 1$, one has*

$$f_{\Delta(n,r)} \leq \min_{x\in\Delta_n} B_r(f)(x), \tag{2.8}$$

$$f_{\Delta(n,r)} - f_{\min,\Delta_n} \leq \min_{x\in\Delta(n,r)} B_r(f)(x) - f_{\min,\Delta_n} \leq \max_{x\in\Delta_n}\{B_r(f)(x) - f(x)\}. \tag{2.9}$$

**Proof.** Note that (2.8) follows from inequality (2.4) and the fact that $\mathbb{E}[f(X)] = B_r(f)(x)$ (by (2.3) and (2.7)). For completeness, we recall the easy argument. Fix $x \in \Delta_n$. By the multinomial theorem, $1 = (\sum_{i=1}^n x_i)^r = \sum_{\alpha\in I(n,r)} \frac{r!}{\alpha!}x^\alpha$. Hence, $B_r(f)(x)$ is a convex combination of the values $f(\frac{\alpha}{r})$ ($\alpha \in I(n,r)$), which implies that $B_r(f)(x) \geq \min_{\alpha\in I(n,r)} f(\frac{\alpha}{r}) = f_{\Delta(n,r)}$.

The left most inequality in (2.9) follows directly from (2.8). To show the right most inequality, let $x^*$ be a global minimizer of $f$ over $\Delta_n$, so that $f(x^*) = f_{\min,\Delta_n}$. Then, $\min_{x\in\Delta_n} B_r(f)(x) - f_{\min,\Delta_n}$ is at most $B_r(f)(x^*) - f_{\min,\Delta_n} = B_r(f)(x^*) - f(x^*)$, which concludes the proof. □

The motivation for using Bernstein approximation to study the quantity $f_{\Delta(n,r)}$ is now clear. Indeed, the Bernstein approximation $B_r(f)$ converges uniformly to $f$ as $r \to \infty$, and the minimum of $B_r(f)$ on $\Delta_n$ is lower bounded by $f_{\Delta(n,r)}$.

Our strategy for upper bounding the range $f_{\Delta(n,r)} - f_{\min,\Delta_n}$ will be to upper bound the (possibly larger) range $\max_{x\in\Delta_n}\{B_r(f)(x) - f(x)\}$ – see Theorems 2.8, 2.11, 2.14 and 2.20. Hence our results can be seen as refinements of the previously known results quoted in Theorem 2.1 above.

The following example shows that all inequalities can be strict in relation (2.9).

**Example 2.5.** *Consider the quadratic polynomial* $f = 2x_1^2 + x_2^2 - 5x_1x_2 \in \mathcal{H}_{2,2}$*. Then,* $B_2(f)(x) = x_1^2 + \frac{1}{2}x_2^2 - \frac{5}{2}x_1x_2 + x_1 + \frac{1}{2}x_2$*. One can easily check that* $f_{\min,\Delta_n} = -\frac{17}{32}$ *(attained at the unique minimizer* $\left(\frac{7}{16}, \frac{9}{16}\right)$*),* $\min_{x\in\Delta_2} B_2(f)(x) = \frac{7}{16}$ *(attained at the unique minimizer* $x = \left(\frac{3}{8}, \frac{5}{8}\right)$*), and* $f_{\Delta(2,2)} = -\frac{1}{2}$ *(attained at the unique minimizer* $\left(\frac{1}{2}, \frac{1}{2}\right)$*). In this example, the polynomial* $f$ *and its Bernstein approximation* $B_2(f)(x)$ *do not have a common minimizer over the standard simplex.*

*Moreover, we note that* $f_{\max,\Delta_n} = 2$ *and* $\max_{x\in\Delta_2}\{B_2(f)(x) - f(x)\} = 1$*, so that we have the following chain of strict inequalities:*

$$\begin{aligned} f_{\Delta(2,2)} - f_{\min,\Delta_n} \left(= \frac{1}{32}\right) &< \min_{x\in\Delta_2} B_2(f)(x) - f_{\min,\Delta_n} \left(= \frac{31}{32}\right) \\ &< \max_{x\in\Delta_2}\{B_2(f)(x) - f(x)\} \ (= 1) \\ &< \frac{1}{2}(f_{\max,\Delta_n} - f_{\min,\Delta_n}) \left(= \frac{81}{64}\right), \end{aligned}$$

*which shows that all the inequalities can be strict in (2.9).*

For any polynomial $f = \sum_{\beta\in I(n,d)} f_\beta x^\beta \in \mathcal{H}_{n,d}$, one can write

$$f = \sum_{\beta\in I(n,d)} f_\beta x^\beta = \sum_{\beta\in I(n,d)} \left(f_\beta \frac{\beta!}{d!}\right) \frac{d!}{\beta!} x^\beta.$$

We call $f_\beta \frac{\beta!}{d!}$ ($\beta \in I(n,d)$) the *Bernstein coefficients* of $f$, since they are the coefficients of the polynomial $f$ when it is expressed in the Bernstein basis

$$\left\{\frac{d!}{\beta!} x^\beta : \beta \in I(n,d)\right\}$$

of $\mathcal{H}_{n,d}$. Using the multinomial theorem (as in the proof of Lemma 2.4), one can see that, for $x \in \Delta_n$, $f(x)$ is a convex combination of its Bernstein coefficients $f_\beta \frac{\beta!}{d!}$ ($\beta \in I(n,d)$). Therefore, one has

$$\min_{\beta\in I(n,d)} f_\beta \frac{\beta!}{d!} \le f_{\min,\Delta_n} \le f(x) \le f_{\max,\Delta_n} \le \max_{\beta\in I(n,d)} f_\beta \frac{\beta!}{d!}. \tag{2.10}$$

We will use the following result of [21], which bounds the range of the Bernstein coefficients in terms of the range of function values.

**Theorem 2.6.** *[21, Theorem 2.2] For any polynomial* $f = \sum_{\beta\in I(n,d)} f_\beta x^\beta \in \mathcal{H}_{n,d}$ *and* $x \in \Delta_n$*, one has*

$$f_{\max,\Delta_n} - f_{\min,\Delta_n} \le \max_{\beta\in I(n,d)} f_\beta \frac{\beta!}{d!} - \min_{\beta\in I(n,d)} f_\beta \frac{\beta!}{d!} \le \binom{2d-1}{d} d^d (f_{\max,\Delta_n} - f_{\min,\Delta_n}).$$

## 2.3 New proofs for the PTAS results

We now give an alternative proof for the PTAS property. More precisely, we show error bounds for four different cases separately, i.e., for the quadratic case (see Corollary 2.9), the cubic case (see Corollary 2.12), the square-free case (see Corollary 2.15), and the general case (see Corollary 2.21). In particular, the error bounds for the first three cases in Corollaries 2.9, 2.12 and 2.15 refine the error bound for the last case in Corollary 2.21.

Recall that we use $\phi_\alpha$ to denote the monomial $x^\alpha$ for $\alpha \in \mathbb{N}^n$, i.e., we set $\phi_\alpha(x) = x^\alpha$.

### 2.3.1 Quadratic polynomial optimization over the standard simplex

We first recall the explicit Bernstein approximation of the monomials of degree at most two, i.e., we compute $B_r(\phi_{\mathbf{e}_i})$, $B_r(\phi_{2\mathbf{e}_i})$ and $B_r(\phi_{\mathbf{e}_i+\mathbf{e}_j})$. We give a proof for clarity.

**Lemma 2.7.** *For $r \geq 1$ one has $B_r(\phi_{\mathbf{e}_i})(x) = x_i, B_r(\phi_{2\mathbf{e}_i})(x) = \frac{1}{r}x_i(1-x_i)+x_i^2$, and $B_r(\phi_{\mathbf{e}_i+\mathbf{e}_j})(x) = \frac{r-1}{r}x_ix_j$ for all $x \in \Delta_n$.*

**Proof.** By the definition (2.7), one has:

$$
\begin{aligned}
B_r(\phi_{\mathbf{e}_i})(x) &= \sum_{\alpha\in I(n,r)} \frac{\alpha_i}{r}\frac{r!}{\alpha!}x^\alpha = x_i \sum_{\beta\in I(n,r-1)} \frac{(r-1)!}{\beta!}x^\beta = x_i(\sum_{i=1}^n x_i)^{r-1} = x_i, \\
B_r(\phi_{\mathbf{e}_i+\mathbf{e}_j})(x) &= \sum_{\alpha\in I(n,r)} \frac{\alpha_i\alpha_j}{r^2}\frac{r!}{\alpha!}x^\alpha = \frac{r-1}{r}x_ix_j \sum_{\beta\in I(n,r-2)} \frac{(r-2)!}{\beta!}x^\beta = \frac{r-1}{r}x_ix_j,
\end{aligned}
$$

$$
\begin{aligned}
B_r(\phi_{2\mathbf{e}_i})(x) &= \sum_{\alpha\in I(n,r)} \frac{\alpha_i^2}{r^2}\frac{r!}{\alpha!}x^\alpha \\
&= \frac{r-1}{r}x_i^2 \sum_{\beta\in I(n,r-2)} \frac{(r-2)!}{\beta!}x^\beta + \frac{1}{r}x_i \sum_{\beta\in I(n,r-1)} \frac{(r-1)!}{\beta!}x^\beta \\
&= \frac{r-1}{r}x_i^2 + \frac{1}{r}x_i = \frac{1}{r}x_i(1-x_i)+x_i^2,
\end{aligned}
$$

where we have used at several places the multinomial theorem (and the fact that an empty summation is equal to 0). $\square$

Consider now a quadratic polynomial $f = x^T Qx \in \mathcal{H}_{n,2}$. By Lemma 2.7, its Bernstein approximation on the standard simplex is given by

$$B_r(f)(x) = \frac{1}{r}\sum_{i=1}^{n} Q_{ii}x_i + (1-\frac{1}{r})f(x) \quad \forall x \in \Delta_n. \tag{2.11}$$

**Theorem 2.8.** *For any polynomial $f = x^T Qx \in \mathcal{H}_{n,2}$ and $r \geq 1$, one has*

$$\max_{x\in\Delta_n}\{B_r(f)(x) - f(x)\} \leq \frac{Q_{\max} - f_{\min,\Delta_n}}{r} \leq \frac{f_{\max,\Delta_n} - f_{\min,\Delta_n}}{r}.$$

*setting $Q_{\max} = \max_{i\in[n]} Q_{ii}$.*

**Proof.** Using (2.11), one obtains that

$$\begin{aligned} rB_r(f)(x) &= \sum_{i=1}^{n} Q_{ii}x_i + (r-1)f(x) \\ &\leq \max_{x\in\Delta_n}\sum_{i=1}^{n} Q_{ii}x_i + rf(x) - \min_{x\in\Delta_n} f(x) \\ &= \max_i Q_{ii} - f_{\min,\Delta_n} + rf(x) \\ &\leq f_{\max,\Delta_n} - f_{\min,\Delta_n} + rf(x), \end{aligned}$$

where in the last inequality we have used the fact that $\max_i Q_{ii} \leq f_{\max,\Delta_n}$, since $Q_{ii} = f(\mathbf{e}_i) \leq f_{\max,\Delta_n}$ for $i \in [n]$. This gives the two right-most inequalities in the theorem. □

Combining Theorem 2.8 with Lemma 2.4, we obtain the following corollary, which gives the PTAS result by Bomze and De Klerk [8, Theorem 3.2].

**Corollary 2.9.** *For any polynomial $f = x^T Qx \in \mathcal{H}_{n,2}$ and $r \geq 1$, one has*

$$f_{\Delta(n,r)} - f_{\min,\Delta_n} \leq \frac{Q_{\max} - f_{\min,\Delta_n}}{r} \leq \frac{f_{\max,\Delta_n} - f_{\min,\Delta_n}}{r}.$$

We note that the proof given here is completely elementary and much simpler than the original one in [8]. Our proof is, however, closely related to another proof by Nesterov [72], as we saw earlier in Section 2.2.2.

**Example 2.10.** *Consider the quadratic polynomial $f = \sum_{i=1}^n x_i^2 \in \mathcal{H}_{n,2}$. As $f$ is convex, it is easy to check that $f_{\min,\Delta_n} = \frac{1}{n}$ (attained at $x = \frac{1}{n}\mathbf{e}$) and $f_{\max,\Delta_n} = 1$ (attained at any standard unit vector).*

*For the computation of $f_{\Delta(n,r)}$, it is convenient to write $r$ as $r = kn + s$, where $k \geq 0$ and $0 \leq s < n$. Then we have*

$$f_{\Delta(n,r)} = \frac{1}{n} + \frac{1}{r^2}\frac{s(n-s)}{n},$$

*which is attained at any point $x \in \Delta(n,r)$ having $n-s$ coordinates equal to $\frac{k}{r}$ and $s$ coordinates equal to $\frac{k+1}{r}$. To see this, pick a minimizer $x \in \Delta(n,r)$. First we claim that $x_i - x_j \leq \frac{1}{r}$ for any $i \neq j \in [n]$. Indeed, assume (say) that $x_2 - x_1 > \frac{1}{r}$. Then define the new point $x' \in \Delta(n,r)$ by $x'_1 = x_1 + \frac{1}{r}$, $x'_2 = x_2 - \frac{1}{r}$ and $x'_i = x_i$ for all $i \neq 1, 2$ and observe that $f(x') < f(x)$, which contradicts the optimality of $x$. Therefore, the coordinates of $x$ can take at most two possible values $\frac{h}{r}, \frac{h+1}{r}$ for some $0 \leq h \leq r-1$ and it is easy to see these two values belong to $\{\frac{k}{r}, \frac{k+1}{r}\}$. Hence we obtain that*

$$f_{\Delta(n,r)} - f_{\min,\Delta_n} = \frac{1}{r^2}\frac{s(n-s)}{n} \quad \textit{and} \quad \frac{f_{\Delta(n,r)} - f_{\min,\Delta_n}}{f_{\max,\Delta_n} - f_{\min,\Delta_n}} = \frac{1}{r^2}\frac{s(n-s)}{n-1}.$$

*We observe that this latter ratio might be in the order $\frac{1}{r}$, thus matching the upper bound in Corollary 2.9 in terms of the dependence of the error bound on $r$. For instance, for $r = \frac{3n}{2}$ (i.e., $k = 1$, $s = \frac{n}{2}$), we have that*

$$f_{\Delta(n,r)} - f_{\min,\Delta_n} = \frac{1}{6r-9}(f_{\max,\Delta_n} - f_{\min,\Delta_n}). \tag{2.12}$$

*Moreover, we have $B_r(f)(x) = \frac{1}{r} + (1 - \frac{1}{r})f(x)$, so that*

$$\min_{x \in \Delta_n} B_r(f)(x) - f_{\min,\Delta_n} = \max_{x \in \Delta_n}\{B_r(f)(x) - f(x)\} = \frac{1}{r}(f_{\max,\Delta_n} - f_{\min,\Delta_n}).$$

*Hence, equality holds throughout in the inequalities of Theorem 2.8, which shows that the upper bound is tight on this example.*

By (2.12) in Example 2.10, there does not exist any $\epsilon > 0$ such that, for any quadratic form $f$,

$$f_{\Delta(n,r)} - f_{\min,\Delta_n} \leq \frac{1}{r^{1+\epsilon}}\left(f_{\max,\Delta_n} - f_{\min,\Delta_n}\right) \quad \forall r \geq 1.$$

Hence, the error bound in Corollary 2.9 is tight in terms of its dependence on $r$. On the other hand, one may easily show that, for the polynomial $f$ in Example 2.10,

$$\rho_r(f) = \frac{f_{\Delta(n,r)} - f_{\min,\Delta_n}}{f_{\max,\Delta_n} - f_{\min,\Delta_n}} \leq \frac{n}{4r^2} = O(1/r^2).$$

Thus, $\limsup_{r\to\infty}(r^2\rho_r(f)) < \infty$, i.e., the asymptotic convergence rate of the sequence $\{\rho_r(f)\}$ for the example is $O(1/r^2)$. It turns out that this will be the case also for the other polynomials considered in Examples 2.13, 2.16 and 2.22 below.

### 2.3.2 Cubic polynomial optimization over the standard simplex

Using similar arguments as for Lemma 2.7, one can compute the Bernstein approximations of the monomials of degree three. Namely, for distinct $i, j, k \in [n]$ and $x \in \Delta_n$,

$$\begin{aligned}
B_r(\phi_{3\mathbf{e}_i})(x) &= \frac{1}{r^2}x_i + \frac{3(r-1)}{r^2}x_i^2 + \frac{(r-1)(r-2)}{r^2}x_i^3, \\
B_r(\phi_{2\mathbf{e}_i+\mathbf{e}_j})(x) &= \frac{(r-1)}{r^2}x_i x_j + \frac{(r-1)(r-2)}{r^2}x_i^2 x_j, \\
B_r(\phi_{\mathbf{e}_i+\mathbf{e}_j+\mathbf{e}_k})(x) &= \frac{(r-1)(r-2)}{r^2}x_i x_j x_k.
\end{aligned}$$

We show the following result.

**Theorem 2.11.** *For any polynomial $f \in \mathcal{H}_{n,3}$ and $r \geq 2$, one has*

$$\max_{x \in \Delta_n}\{B_r(f)(x) - f(x)\} \leq \left(\frac{4}{r} - \frac{4}{r^2}\right)(f_{\max,\Delta_n} - f_{\min,\Delta_n}).$$

**Proof.** Consider a cubic polynomial $f \in \mathcal{H}_{n,3}$ of the form

$$f = \sum_{i=1}^n f_i x_i^3 + \sum_{1 \leq i < j \leq n}(f_{ij}x_i x_j^2 + g_{ij}x_i^2 x_j) + \sum_{1 \leq i < j < k \leq n} f_{ijk}x_i x_j x_k.$$

Applying the above description for the Bernstein approximation of degree 3 monomials, the Bernstein approximation of $f$ at any $x \in \Delta_n$ reads

$$\begin{aligned}
B_r(f)(x) &= \frac{(r-1)(r-2)}{r^2}f(x) \\
&+ \frac{1}{r^2}\left[\sum_{i=1}^n f_i x_i + (r-1)\left(\sum_{i=1}^n 3f_i x_i^2 + \sum_{i<j}(f_{ij}+g_{ij})x_i x_j\right)\right]. \qquad (2.13)
\end{aligned}$$

Evaluating $f$ at $\mathbf{e}_i$ and at $(\mathbf{e}_i + \mathbf{e}_j)/2$ yields, respectively, the relations:

$$f_{\min,\Delta_n} \leq f_i \leq f_{\max,\Delta_n}, \qquad (2.14)$$

$$f_i + f_j + f_{ij} + g_{ij} \leq 8 f_{\max,\Delta_n}. \qquad (2.15)$$

Using (2.15) and the fact that $\sum_{i=1}^n x_i = 1$, one can obtain

$$\sum_{i<j}(f_{ij}+g_{ij})x_ix_j \le \sum_{i<j}(8f_{\max,\Delta_n}-f_i-f_j)x_ix_j = 8f_{\max,\Delta_n}\sum_{i<j}x_ix_j - \sum_{i=1}^n f_ix_i(1-x_i). \tag{2.16}$$

Combining (2.13) and (2.16), one obtains that, for any $x \in \Delta_n$,

$$\begin{aligned} r^2B_r(f)(x) &= (r-1)(r-2)f(x) + \sum_{i=1}^n f_ix_i \\ &+ (r-1)\left(\sum_{i=1}^n 3f_ix_i^2 + \sum_{i<j}(f_{ij}+g_{ij})x_ix_j\right) \\ &\le (r-1)(r-2)f(x) - (r-2)\sum_{i=1}^n f_ix_i \\ &\quad +(r-1)\left(\sum_{i=1}^n 4f_ix_i^2 + 8f_{\max,\Delta_n}\sum_{i<j}x_ix_j\right). \end{aligned}$$

We now use (2.14) to bound the two inner summations as follows:

$$\begin{aligned} -\sum_i f_ix_i &\le -f_{\min,\Delta_n}\sum_i x_i = -f_{\min,\Delta_n} \\ \text{and } \sum_{i=1}^n 4f_ix_i^2 &+ 8f_{\max,\Delta_n}\sum_{i<j}x_ix_j \le 4f_{\max,\Delta_n}(\sum_i x_i)^2 = 4f_{\max,\Delta_n}. \end{aligned}$$

This implies:

$$\begin{aligned} r^2(B_r(f)(x)-f(x)) &\le -(3r-2)f_{\min,\Delta_n} - (r-2)f_{\min,\Delta_n} + 4(r-1)f_{\max,\Delta_n} \\ &= 4(r-1)(f_{\max,\Delta_n}-f_{\min,\Delta_n}), \end{aligned}$$

which concludes the proof. □

Combining Theorem 2.11 with Lemma 2.4, we obtain the following error bound.

**Corollary 2.12.** *For any polynomial $f \in \mathcal{H}_{n,3}$ and $r \ge 2$, one has*

$$f_{\Delta(n,r)} - f_{\min,\Delta_n} \le \left(\frac{4}{r}-\frac{4}{r^2}\right)(f_{\max,\Delta_n}-f_{\min,\Delta_n}).$$

This result is a bit stronger than the result by De Klerk et al. [21, Theorem 3.3], which states that $f_{\Delta(n,r)} - f_{\min,\Delta_n} \leq \frac{4}{r}(f_{\max,\Delta_n} - f_{\min,\Delta_n})$.

**Example 2.13.** *Consider the cubic polynomial $f = x_1^3 + x_2^3 \in \mathcal{H}_{2,3}$. One can check that $f_{\max,\Delta_n} = 1, f_{\min,\Delta_n} = \frac{1}{4}$,*

$$f_{\Delta(2,r)} = \begin{cases} 1/4 & \text{if } r \text{ is even,} \\ \frac{1}{4} + \frac{3}{4r^2} & \text{if } r \text{ is odd.} \end{cases}$$

*Moreover, one can check that $B_r(f)(x) = 1 + \left(\frac{3}{r} - 3\right) x_1 x_2$ and*

$$\min_{x \in \Delta_2} B_r(f)(x) = \frac{1}{4} + \frac{3}{4r}.$$

*Hence, for any integer $r \geq 2$, one has strict inequality*

$$\min_{x \in \Delta_2} B_r(f)(x) > f_{\Delta(2,r)}.$$

*Moreover, for $r \geq 2$,*

$$\begin{aligned} \min_{x \in \Delta_2} B_r(f)(x) - f_{\min,\Delta_n} &= \max_{x \in \Delta_2} \{B_r(f)(x) - f(x)\} = \frac{3}{4r} = \frac{1}{r}(f_{\max,\Delta_n} - f_{\min,\Delta_n}) \\ &< (\frac{4}{r} - \frac{4}{r^2})(f_{\max,\Delta_n} - f_{\min,\Delta_n}). \end{aligned}$$

*On the other hand, for odd $r$, the range $f_{\Delta(2,r)} - f_{\min,\Delta_n}$ is equal to*

$$\frac{3}{4r^2} = \frac{1}{r^2}(f_{\max,\Delta_n} - f_{\min,\Delta_n})$$

*and thus grows proportionally to $\frac{1}{r^2}$.*

### 2.3.3 Square-free polynomial optimization over the standard simplex

Here we consider square-free (or multilinear) polynomials, involving only monomials $x^I = \prod_{i \in I} x_i$ for $I \subseteq [n]$. The Bernstein approximation of the square-free monomial $\phi_{\mathbf{e}_I}(x) = x^I$, with $d = |I|$, is given by

$$B_r(\phi_{\mathbf{e}_I})(x) = \sum_{\alpha \in I(n,r)} \frac{\alpha^I}{r^d} \frac{r!}{\alpha!} x^\alpha = \frac{r^{\underline{d}}}{r^d} x^I \sum_{\alpha \in I(n,r-d)} \frac{(r-d)!}{\alpha!} x^\alpha = \frac{r^{\underline{d}}}{r^d} x^I (\sum_i x_i)^{r-d} = \frac{r^{\underline{d}}}{r^d} x^I$$

for $x \in \Delta_n$. Recall that, for an integer $r \geq 1$, $r^{\underline{d}} = r(r-1)\cdots(r-d+1)$ and $r^{\underline{d}} = 0$ if $r < d$. Hence the Bernstein approximation of the square-free polynomial $f = \sum_{I \subseteq [n], |I|=d} f_I x^I$ satisfies

$$B_r(f)(x) = \frac{r^{\underline{d}}}{r^d} f(x) \quad \forall x \in \Delta_n,$$

which implies the following identities:

$$\min_{x \in \Delta_n} B_r(f)(x) - \underline{f} = \max_{x \in \Delta_n} \{B_r(f)(x) - f(x)\} = -\left(1 - \frac{r^{\underline{d}}}{r^d}\right) f_{\min,\Delta_n}. \tag{2.17}$$

**Theorem 2.14.** *For any square-free polynomial $f \in \mathcal{H}_{n,d}$ and $r \geq 1$, one has*

$$\max_{x \in \Delta_n} \{B_r(f)(x) - f(x)\} \leq \left(1 - \frac{r^{\underline{d}}}{r^d}\right) (f_{\max,\Delta_n} - f_{\min,\Delta_n}) \quad \forall x \in \Delta_n.$$

**Proof.** We use (2.17). For degree $d = 1$ the result is clear and, for $d \geq 2$, we use the fact that $f_{\max,\Delta_n} \geq 0$ since $f(\mathbf{e}_i) = 0$ for any $i \in [n]$. □

Combining with Lemma 2.4 we obtain the following error bound.

**Corollary 2.15.** *For any square-free polynomial $f \in \mathcal{H}_{n,d}$ and $r \geq 1$, one has*

$$f_{\Delta(n,r)} - f_{\min,\Delta_n} \leq \left(1 - \frac{r^{\underline{d}}}{r^d}\right) (f_{\max,\Delta_n} - f_{\min,\Delta_n}).$$

This result was first shown by Nesterov [72, Theorem 2] (see also De Klerk et al. [21, Remark 3.4]). In fact, our proof is again closely related to the one by Nesterov [72], as we saw earlier in Section 2.2.2.

**Example 2.16.** *Consider the square-free polynomial $f = -x_1x_2$. Then, $B_r(f)(x) = -\frac{r-1}{r}x_1x_2$ and one can check that $f_{\max,\Delta_n} = 0$, $f_{\min,\Delta_n} = -\frac{1}{4}$, and $\min_{x \in \Delta_2} B_r(f)(x) = -\frac{1}{4}\frac{r-1}{r}$. Moreover,*

$$f_{\Delta(2,r)} = \begin{cases} -\frac{1}{4} & \text{if } r \text{ is even}, \\ -\frac{1}{4} + \frac{1}{4r^2} & \text{if } r \text{ is odd}. \end{cases}$$

*Hence, for any integer $r \geq 2$, one has strict inequality: $\min_{x \in \Delta_2} B_r(f)(x) > f_{\Delta(n,r)}$. Moreover, as $\min_{x \in \Delta_2} B_r(f)(x) - f_{\min,\Delta_n} = \max_{x \in \Delta_2}\{B_r(f)(x) - f(x)\} = \frac{1}{4r} = \frac{1}{r}(f_{\max,\Delta_n} - f_{\min,\Delta_n})$, the upper bound from Theorem 2.14 is tight on this example. On the other hand, $f_{\Delta(2,r)} - f_{\min,\Delta_n} = \frac{1}{4r^2} = \frac{1}{r^2}(f_{\max,\Delta_n} - f_{\min,\Delta_n})$ for odd $r$, and thus the range $f_{\Delta(2,r)} - f_{\min,\Delta_n}$ grows proportionally to $\frac{1}{r^2}$.*

### 2.3.4 General polynomial optimization over the standard simplex

We now deal with the minimization of an arbitrary polynomial $f \in \mathcal{H}_{n,d}$. In order to be able to bound the minimum of $B_r(f)$ over $\Delta_n$ we need an explicit description of the Bernstein approximation of $f$.

#### Bernstein approximation over the standard simplex of an arbitrary monomial

Here we work out an explicit description of the Bernstein approximation of arbitrary monomials $\phi_\beta(x) = x^\beta$ ($\beta \in I(n,d)$). The key ingredient is to express it in terms of the moments of the multinomial distribution.

Fix $x = (x_1, \dots, x_n) \in \Delta_n$ and assume $Y = (Y_1, \dots, Y_n)$ has the multinomial distribution with parameters $r$, $n$ and $x_1, \dots, x_n$ (where $x \in \Delta_n$) as in Section 2.2.1. Then, given $\alpha \in I(n,r)$, by (2.2), the probability of the event $Y = \alpha$ is equal to $\frac{r!}{\alpha!}x^\alpha$. Therefore, for $\beta \in \mathbb{N}^n$, the $\beta$-th moment of this multinomial distribution is given by

$$m^\beta_{(n,r)} := \mathbb{E}[Y^\beta] = \sum_{\alpha \in I(n,r)} \alpha^\beta \frac{r!}{\alpha!} x^\alpha. \tag{2.18}$$

Comparing with the definition of the Bernstein approximation of $\phi_\beta(x) = x^\beta$ we find the identity

$$B_r(\phi_\beta)(x) = \sum_{\alpha \in I(n,r)} (\frac{\alpha}{r})^\beta \frac{r!}{\alpha!} x^\beta = \frac{1}{r^{|\beta|}} m^\beta_{(n,r)}.$$

Combining [40, relation (34.18)] and [40, relation (35.5)], we can obtain an alternative formula for the moments $m^\beta_{(n,r)}$ of the multinomial distribution in terms of *Stirling numbers of the second kind*, see Theorem 2.18 below.

**Definition 2.17.** *For integers $n, k \in \mathbb{N}$, the Stirling number of the second kind $S(n,k)$ counts the number of ways of partitioning a set of $n$ objects into $k$ nonempty subsets. Thus $S(n,k) = 0$ if $k > n$, $S(n,0) = 0$ if $n \geq 1$, and $S(0,0) = 1$ by convention.*

*For more information about the Stirling number of the second kind, see Appendix A.*

**Theorem 2.18.** *[40] For $\beta \in \mathbb{N}^n$, one has*

$$m^\beta_{(n,r)} = \sum_{\alpha \in \mathbb{N}^n : \alpha \leq \beta} r^{\underline{|\alpha|}} x^\alpha \prod_{i=1}^n S(\beta_i, \alpha_i).$$

Therefore, we can deduce the explicit formula of the Bernstein approximation for any monomial.

**Corollary 2.19.** *For any monomial $\phi_\beta(x) = x^\beta$, one has*

$$B_r(\phi_\beta)(x) = \frac{1}{r^{|\beta|}} \sum_{\alpha\in\mathbb{N}^n:\alpha\le\beta} r^{\underline{|\alpha|}} x^\alpha \prod_{i=1}^n S(\beta_i,\alpha_i) \quad \forall x \in \Delta_n.$$

For completeness, we will give a self-contained proof for Theorem 2.18 in Appendix B.

### Error bound analysis

We show the following error bound for the Bernstein approximation of order $r$ of an arbitrary polynomial on the standard simplex.

**Theorem 2.20.** *For any polynomial $f \in \mathcal{H}_{n,d}$ and $r \ge 1$, one has*

$$\begin{aligned}\max\nolimits_{x\in\Delta_n}\{B_r(f)(x) - f(x)\} &\le \left(1 - \tfrac{r^{\underline{d}}}{r^d}\right)\left(\max_{\beta\in I(n,d)} f_\beta \frac{\beta!}{d!} - \min_{\beta\in I(n,d)} f_\beta \frac{\beta!}{d!}\right) \\ &\le \left(1 - \tfrac{r^{\underline{d}}}{r^d}\right)\tbinom{2d-1}{d} d^d (f_{\max,\Delta_n} - f_{\min,\Delta_n}).\end{aligned}$$

**Proof.** Consider a polynomial $f = \sum_{\beta\in I(n,d)} f_\beta x^\beta \in \mathcal{H}_{n,d}$ and $x \in \Delta_n$. Applying Corollary 2.19, we can write the Bernstein approximation of $f$ at $x \in \Delta_n$ as follows:

$$B_r(f)(x) = \frac{1}{r^d} \sum_{\beta\in I(n,d)} f_\beta \sum_{\alpha:0\le\alpha\le\beta} r^{\underline{|\alpha|}} x^\alpha \prod_{i=1}^n S(\beta_i,\alpha_i).$$

Therefore,

$$r^d B_r(f)(x) = r^{\underline{d}} f(x) + \sum_{\beta\in I(n,d)} f_\beta \sum_{\alpha:0\le\alpha\le\beta,\alpha\ne\beta} r^{\underline{|\alpha|}} x^\alpha \prod_{i=1}^n S(\beta_i,\alpha_i),$$

and thus

$$r^d(B_r(f)(x) - f(x)) = -(r^d - r^{\underline{d}}) f(x) + \sum_{\beta\in I(n,d)} f_\beta \sum_{\alpha:0\le\alpha\le\beta,\alpha\ne\beta} r^{\underline{|\alpha|}} x^\alpha \prod_{i=1}^n S(\beta_i,\alpha_i).$$

Using (2.10), we have $f(x) \geq \min_{\beta \in I(n,d)} f_\beta \frac{\beta!}{d!}$ and $f_\beta \frac{\beta!}{d!} \leq \max_{\beta' \in I(n,d)} f_{\beta'} \frac{\beta'!}{d!}$, which permits to derive the following inequality:

$$r^d(B_r(f)(x) - f(x)) \leq$$
$$-(r^d - r^{\underline{d}}) \min_{\beta \in I(n,d)} f_\beta \frac{\beta!}{d!} + \max_{\beta \in I(n,d)} f_\beta \frac{\beta!}{d!} \left( \underbrace{\sum_{\beta \in I(n,d)} \frac{d!}{\beta!} \sum_{\alpha: 0 \leq \alpha \leq \beta, \alpha \neq \beta} x^\alpha r^{\underline{|\alpha|}} \prod_{i=1}^n S(\beta_i, \alpha_i)}_{\sigma} \right). \tag{2.19}$$

It now suffices to upper bound the right handside of the inequality (2.19) and to show that the inner summation $\sigma$ is equal to $r^d - r^{\underline{d}}$. Indeed,

$$\begin{aligned}
\sigma &= \sum_{\beta \in I(n,d)} \frac{d!}{\beta!} \sum_{\alpha: 0 \leq \alpha \leq \beta, \alpha \neq \beta} x^\alpha r^{\underline{|\alpha|}} \prod_{i=1}^n S(\beta_i, \alpha_i) \\
&= \sum_{\alpha \in \mathbb{N}^n} x^\alpha r^{\underline{|\alpha|}} \sum_{\beta \in I(n,d): \beta \geq \alpha, \beta \neq \alpha} \frac{d!}{\beta!} \prod_{i=1}^n S(\beta_i, \alpha_i) \\
&= \sum_{k=1}^{d-1} \sum_{\alpha \in I(n,k)} x^\alpha r^{\underline{|\alpha|}} \left( \sum_{\beta \in I(n,d): \beta \geq \alpha} \frac{d!}{\beta!} \prod_{i=1}^n S(\beta_i, \alpha_i) \right) \\
&= \sum_{k=1}^{d-1} r^{\underline{k}} \sum_{\alpha \in I(n,k)} x^\alpha \left( \frac{k!}{\alpha!} S(d,k) \right) && \text{[using Lemma A.2]} \\
&= \sum_{k=1}^{d-1} r^{\underline{k}} S(d,k) (\sum_i x_i)^k = \sum_{k=1}^{d-1} r^{\underline{k}} S(d,k) = r^d - r^{\underline{d}}. && \text{[using Lemma A.3]}
\end{aligned}$$

Using this identity for the summation $\sigma$ in the inequality (2.19) we obtain

$$\begin{aligned}
r^d \left( \min_{x \in \Delta_n} B_r(f)(x) - f_{\min,\Delta_n} \right) &\leq r^d \max_{x \in \Delta_n} \{ B_r(f)(x) - f(x) \} \\
&\leq (r^d - r^{\underline{d}}) \left( \max_{\beta \in I(n,d)} f_\beta \frac{\beta!}{d!} - \min_{\beta \in I(n,d)} f_\beta \frac{\beta!}{d!} \right).
\end{aligned}$$

By combining with Theorem 2.6 we obtain the claimed inequalities of Theorem 2.20 and this concludes the proof. □

Combining Theorem 2.20 with Lemma 2.4, we obtain the following error bound, which was first shown in [21, Theorem 1.3].

**Corollary 2.21.** *For any polynomial $f \in \mathcal{H}_{n,d}$ and $r \geq 1$, one has*

$$f_{\Delta(n,r)} - f_{\min,\Delta_n} \leq \left(1 - \frac{r^{\underline{d}}}{r^d}\right) \binom{2d-1}{d} d^d (f_{\max,\Delta_n} - f_{\min,\Delta_n}).$$

**Example 2.22.** *We consider here the problem of minimizing the polynomial $f = \sum_{i=1}^n x_i^d$ ($n \geq 2$) over the standard simplex for any degree $d \geq 2$, thus extending the case $d = 2$ considered in Example 2.10 and the case $d = 3, n = 2$ considered in Example 2.13. As $f$ is convex on $\mathbb{R}^n_+$ it follows that $f_{\max,\Delta_n} = 1$ and $f_{\min,\Delta_n} = \frac{1}{n^{d-1}}$. We now compute the minimum over the regular grid $\Delta(n,r)$. As in Example 2.10 set $r = kn + s$ where $k, s \in \mathbb{N}$ with $s \leq n-1$. We show that $f_{\Delta(n,r)}$ is attained at any point $x$ having $s$ components equal to $\frac{k+1}{r}$ and $n-s$ components equal to $\frac{k}{r}$, so that*

$$f_{\Delta(n,r)} = s\left(\frac{k+1}{r}\right)^d + (n-s)\left(\frac{k}{r}\right)^d.$$

*For this pick a minimizer $x$ of $f$ over $\Delta(n,r)$ and it suffices to show that $x_i - x_j \leq \frac{1}{r}$ for all $i, j \in [n]$. If (say) $x_2 - x_1 > \frac{1}{r}$ then we claim that*

$$f(x_1 + \frac{1}{r}, x_2 - \frac{1}{r}, x_3, \ldots, x_n) < f(x_1, x_2, x_3, \ldots, x_n),$$

*which contradicts the minimality assumption on $x$. One can see the above claim as follows: set $\sigma = 1 - \sum_{i=3}^n x_i$, consider the function $\phi(t) = t^d + (\sigma - t)^d$ for $t$ satisfying $0 \leq t < \frac{1}{2}(\sigma - \frac{1}{r})$, and verify (using elementary calculus) that $\phi(t + \frac{1}{r}) < \phi(t)$ for any such $t$. Therefore, we have*

$$\begin{aligned}
f_{\Delta(n,r)} - f_{\min,\Delta_n} &= (n-s)\left(\frac{k}{r}\right)^d + s\left(\frac{k+1}{r}\right)^d - \frac{n}{n^d} \\
&= (n-s)\left(\left(\frac{k}{r}\right)^d - \frac{1}{n^d}\right) + s\left(\left(\frac{k+1}{r}\right)^d - \frac{1}{n^d}\right) \\
&= \frac{n-s}{n^d}\left(\left(1 - \frac{s}{r}\right)^d - 1\right) + \frac{s}{n^d}\left(\left(1 - \frac{s-n}{r}\right)^d - 1\right) \\
&= \frac{s(n-s)}{n^d} \sum_{i=2}^d \binom{d}{i} \frac{(n-s)^{i-1} + (-1)^i s^{i-1}}{r^i}.
\end{aligned}$$

*Using the fact that $s, n-s \leq n$, for any $r \geq n$ we can bound the above summation as follows:*

$$\begin{aligned}\sum_{i=2}^{d}\binom{d}{i}\frac{(n-s)^{i-1}+(-1)^i s^{i-1}}{r^i} &\leq \frac{2}{n}\sum_{i=2}^{d}\binom{d}{i}\left(\frac{n}{r}\right)^i \leq \frac{2}{n}\left(\frac{n}{r}\right)^2\sum_{i=2}^{d}\binom{d}{i}\\ &\leq \frac{2}{n}\left(\frac{n}{r}\right)^2 2^d = \frac{2^{d+1}n}{r^2}.\end{aligned}$$

*Combining with the bound $s(n-s) \leq \frac{n^2}{4}$, we deduce that*

$$f_{\Delta(n,r)} - f_{\min,\Delta_n} \leq \frac{n^2}{4}\frac{1}{n^d}\frac{2^{d+1}n}{r^2} = \frac{2^{d-1}}{n^{d-3}r^2}.$$

*Therefore, for any $r \geq n \geq 2$ and $d \geq 3$, one has*

$$\frac{f_{\Delta(n,r)} - f_{\min,\Delta_n}}{f_{\max,\Delta_n} - f_{\min,\Delta_n}} \leq \frac{2^{d-1}}{n^{d-3}r^2}\frac{n^{d-1}}{n^{d-1}-1} = \frac{2^{d-1}}{r^2}\frac{n^2}{n^{d-1}-1} \leq \frac{2^d}{r^2}. \tag{2.20}$$

*Hence we see that for any degree $d \geq 3$ the ratio is in the order $\frac{1}{r^2}$. Recall that for degree $d = 2$ it was observed in Example 2.10 that it can be in the order $\frac{1}{r}$ for certain values of $r$ (e.g., for $r = \frac{3n}{2}$).*

## 2.4 Concluding remarks

A final comment concerns the convergence ratio $\rho_r(f)$ defined as in (1.17):

$$\rho_r(f) = \frac{f_{\Delta(n,r)} - f_{\min,\Delta_n}}{f_{\max,\Delta_n} - f_{\min,\Delta_n}} \quad r = 1, 2, \ldots$$

for a given polynomial $f \in \mathcal{H}_{n,d}$.

On the one hand, our results in Corollaries 2.9, 2.12, 2.15, and 2.21 imply that for any homogeneous polynomial $f$ of degree $d$,

$$\rho_r(f) \leq \frac{C(d)}{r} = O(1/r),$$

where $C(d)$ is a constant depending only on $d$.

On the other hand, in all the examples presented in this chapter, a sharper estimate

$$\rho_r(f) \leq \frac{C(f)}{r^2} = O(1/r^2)$$

holds, where $C(f)$ is a constant depending on $f$; see Examples 2.10, 2.13, 2.16 and 2.22. Therefore, it remains an open problem to determine the exact dependence on $r$ of the convergence ratio $\rho_r(f)$ in general. We will give a partial answer in the following Chapter 3.

# Chapter 3

# A refined error analysis

## 3.1 Introduction

In this chapter, we study the parameter $f_{\Delta(n,r)}$, which was defined in (1.9):

$$f_{\Delta(n,r)} = \min \ f(x) \text{ s.t. } x \in \Delta(n,r) = \{x \in \Delta_n : rx \in \mathbb{N}^n\}.$$

In particular, we consider the question posed in Section 2.4 concerning the exact dependence on $r$ of the convergence ratio:

$$\rho_r(f) = \frac{f_{\Delta(n,r)} - f_{\min,\Delta_n}}{f_{\max,\Delta_n} - f_{\min,\Delta_n}} \quad r = 1, 2, \ldots$$

By Theorem 2.1, we know that $\rho_r(f)$ is of the order $O(1/r)$. On the other hand, in Chapter 2, several examples are given where $\rho_r(f)$ is in fact of the order $O(1/r^2)$ and the question is posed whether this can be true in general.

Here, we give a partial answer to this question. For any polynomial $f \in \mathcal{H}_{n,d}$, we will show that, under some conditions on $f$,

$$\rho_r(f) \leq \frac{C(f)}{r^2} = O\left(\frac{1}{r^2}\right),$$

where the constant $C(f)$ depends on the polynomial $f$. More precisely, for any quadratic $f$, we show that $\rho_r(f) \leq m/r^2$ if $f$ has a global minimizer with denominator $m$ (see Theorem 3.7). In view of Vavasis' result [95] on the existence of rational minimizers for quadratic programming, this implies that $\rho_r(f) = O(1/r^2)$ for any quadratic $f$ (see Corollary 3.9). For polynomials $f$ of degree $d \geq 3$, when $f$ admits a rational global minimizer, we also show that $\rho_r(f) = O(1/r^2)$ (see Corollaries 3.13 and 3.16).

As explained in Section 2.2.1, our approach in Chapter 2 can be put in the more general context of the framework introduced by Lasserre [47, 53], based on reformulating any polynomial optimization problem as an optimization problem over measures. When applied to our setting, this implies the following upper bound:

$$f_{\Delta(n,r)} \leq \mathbb{E}_\mu(f) = \int_{\Delta(n,r)} f(x)\mu(dx)$$

for any probability measure $\mu$ on $\Delta(n,r)$. So the work in Chapter 2 is based on selecting the multinomial distribution with appropriate parameters as measure $\mu$.

In this chapter we will replace the multinomial distribution by the hypergeometric distribution, and we therefore review some necessary background material next.

## 3.2 The multivariate hypergeometric distribution

Consider a box containing $m$ balls of $n$ colors, of which $m_i$ are of color $i$ for $i = 1, \ldots, n$. Thus $\sum_{i=1}^n m_i = m$. We draw $r$ balls randomly from the box without replacement. This defines the random variable $Y_i$ as the number of balls of color $i$ in a random sample of $r$ balls.

Then, $Y = (Y_1, \ldots, Y_n)$ has the *multivariate hypergeometric distribution*, with parameters $r$, $n$ and $m_1, \ldots, m_n$ (with $\sum_{i=1}^n m_i = m$). Given $\alpha \in \mathbb{N}^n$ with $\sum_{i=1}^n \alpha_i = r$, the probability of obtaining the outcome $Y = \alpha$, i.e., with $\alpha_i$ balls of color $i$, is equal to

$$\mathbf{Pr}\left[Y_1 = \alpha_1, \ldots, Y_n = \alpha_n\right] = \frac{\prod_{i=1}^n \binom{m_i}{\alpha_i}}{\binom{m}{r}}. \tag{3.1}$$

Define the random variables

$$X = (X_1, \ldots, X_n) \;\text{ where } X_i = Y_i/r \;\; (i = 1, \ldots, n). \tag{3.2}$$

Thus $X$ takes its values in $\Delta(n,r)$.

As mentioned before, in this chapter we will apply the multivariate hypergeometric distribution to upper bound $f_{\Delta(n,r)}$. More precisely, we will use the following lemma, which is analogous to Lemma 2.4 in Chapter 2.

**Lemma 3.1.** *Let $f = \sum_{\beta \in I(n,d)} f_\beta x^\beta \in \mathcal{H}_{n,d}$ and let $X = (X_1, X_2, \ldots, X_n)$ be as in (3.2). Then, one has*

$$f_{\Delta(n,r)} \leq \mathbb{E}\left[f(X)\right]$$

*and the above inequality can be strict.*

**Proof.** By definition (3.2), the random variable $X$ takes its values in $\Delta(n,r)$, which implies directly that the expected value of $f(X)$ is at least the minimum of $f$ over $\Delta(n,r)$.

In order to show that the inequality can be strict, we consider the following example: $f = 2x_1^2 + x_2^2 - 5x_1x_2$. One has $f_{\min,\Delta_n} = -\frac{17}{32}$ attained at the unique minimizer $(\frac{7}{16}, \frac{9}{16})$. Then we let $m = 16$, $m_1 = 7$ and $m_2 = 9$. When $r = 2$, one can easily check that $f_{\Delta(2,2)} = -\frac{1}{2}$ (attained at the unique minimizer $(\frac{1}{2}, \frac{1}{2})$). On the other hand, when $r = 2$, $\mathbb{E}\left[f(X)\right] = \frac{31}{80}$, and thus $\mathbb{E}\left[f(X)\right] > f_{\Delta(2,2)}$. $\square$

**Remark 3.2.** *To motivate the choice of the hypergeometric distribution over the multinomial distribution, consider the case where $f$ has a rational minimizer $x^*$ in $\Delta(n,m)$, i.e., each component of $x^*$ has denominator $m$.*

*If we now define the random variable $X$ as in (3.1) and (3.2) with $m_i = mx_i^*$ ($i \in [n]$) and $r \leq m$, then*

$$\mathbb{E}\left[f(X)\right] = \sum_{\alpha \in r\Delta(n,r)} \prod_{i=1}^{n} \frac{1}{\binom{m}{r}} \binom{mx_i^*}{\alpha_i} f\left(\frac{\alpha}{r}\right) =: H_r(f)(x^*).$$

*Note that $H_r(f)(x^*)$ is the analog of the Bernstein approximation $B_r(f)(x^*)$ in (2.3). If $r = m$, then the only possible value that $X$ can take is $x^*$. In other words, $H_m(f)(x^*) = f(x^*) = f_{\min,\Delta_n}$, which means finite convergence of $H_r(f)(x^*)$ ($r = 1, 2, \ldots$) to $f_{\min,\Delta_n}$, whereas the convergence $\lim_{r\to\infty} B_r(f)(x^*) = f_{\min,\Delta_n}$ is not finite in general. For instance, consider the polynomial $f = x_1^2 + x_1^2$. It is easy to see $f_{\min,\Delta_2} = \frac{1}{2}$ when $x^* = \left(\frac{1}{2}, \frac{1}{2}\right)$. Then, by Lemma 2.7, one has $B_r(f)(x^*) = \frac{1}{2} + \frac{1}{2r}$, and thus $B_r(f)(x^*) > f_{\min,\Delta_n}$ for any $r \geq 1$.*

In order to apply Lemma 3.1, we need an explicit formula for the $\beta$-th moment of $X$ for any $\beta \in \mathbb{N}^n$:

$$\mathbb{E}\left[X^\beta\right] = \mathbb{E}\left[\prod_{i=1}^{n} X_i{}^{\beta_i}\right].$$

In what follows, we show an explicit formula for $\mathbb{E}\left[X^\beta\right]$ in Corollary 3.4, which is analogous to the explicit formula for the $\beta$-th the moment of the multinomial distribution in Theorem 2.18.

By (3.1), for any $\beta \in \mathbb{N}^n$, the $\beta$-th moment of the multivariate hypergeometric distribution reads as

$$\mathbb{E}\left[Y^\beta\right] = \mathbb{E}\left[\prod_{i=1}^n Y_i^{\beta_i}\right] = \sum_{\alpha \in I(n,r)} \alpha^\beta \frac{\prod_{i=1}^n \binom{m_i}{\alpha_i}}{\binom{m}{r}}.$$

Combining [40, relation (34.18)] and [40, relation (39.6)], we can obtain an alternative formula for $\mathbb{E}\left[Y^\beta\right]$ in terms of the Stirling numbers of the second kind (recall Definition 2.17).

**Theorem 3.3.** *For $\beta \in \mathbb{N}^n$, one has*

$$\mathbb{E}\left[Y^\beta\right] = \sum_{\alpha \in \mathbb{N}^n : \alpha \le \beta} \frac{r^{\underline{|\alpha|}}}{m^{\underline{|\alpha|}}} \prod_{i=1}^n m_i^{\underline{\alpha_i}} S(\beta_i, \alpha_i),$$

*where $r^{\underline{d}} = r(r-1)\cdots(r-d+1)$.*

By Theorem 3.3 one can easily obtain the following explicit formula for $\mathbb{E}\left[X^\beta\right]$, where $X$ is the random variable defined as in (3.2).

**Corollary 3.4.** *For $\beta \in \mathbb{N}^n$, one has*

$$\mathbb{E}\left[X^\beta\right] = \frac{1}{r^{|\beta|}} \sum_{\alpha \in \mathbb{N}^n : \alpha \le \beta} \frac{r^{\underline{|\alpha|}}}{m^{\underline{|\alpha|}}} \prod_{i=1}^n m_i^{\underline{\alpha_i}} S(\beta_i, \alpha_i).$$

## 3.3 The convergence analysis

In this section, we show some new error bounds for the parameter $f_{\Delta(n,r)}$ based on using the multivariate hypergeometric distribution. Similarly as in Chapter 2, we consider four different cases separately, i.e., the quadratic case, the cubic case, the square-free case, and the general case.

On the one hand, the proofs in this section apply similar ideas as the proofs in Chapter 2. On the other hand, due to the difference between the moments of the hypergeometric distribution (in Corollary 3.4) and of the multinomial distribution (in Theorem 2.18), the proofs in this section are more technical.

### 3.3.1 The quadratic case

We now consider the case when $f$ is assumed to be quadratic. First we show the following result.

**Theorem 3.5.** *Let $f = x^TQx \in \mathcal{H}_{n,2}$. For any integers $r$ and $m \geq 2$ such that $1 \leq r \leq m$, one has*

$$f_{\Delta(n,r)} - f_{\Delta(n,m)} \leq \frac{m-r}{r(m-1)} \left(f_{\max,\Delta_n} - f_{\min,\Delta_n}\right).$$

**Proof.** Let $m \geq 2$ and let $x^* \in \Delta(n,m)$ be a minimizer of $f$ over $\Delta(n,m)$, i.e., $f(x^*) = f_{\Delta(n,m)}$, and set $m_i = mx_i^*$ for $i \in [n]$. If $m = 1$, then $r = 1$ and the result is trivial. Now assume $m \geq 2$. Consider the random variable $X = (X_1, \dots, X_n)$ defined as in (3.1) and (3.2). By Corollary 3.4, one has

$$\begin{aligned}
\mathbb{E}[X_i^2] &= \left(\frac{m_i}{m}\right)^2 \left(1 - \frac{m-r}{r(m-1)} + \frac{m(m-r)}{rm_i(m-1)}\right) \quad (i \in [n]),\\
\mathbb{E}[X_iX_j] &= \frac{m_i}{m}\frac{m_j}{m}\left(1 - \frac{m-r}{r(m-1)}\right) \quad (i \neq j \in [n]).
\end{aligned}$$

Then, we have

$$\begin{aligned}
\mathbb{E}\left[f(X)\right] &= \sum_{i,j\in[n]:i\neq j} Q_{ij}\mathbb{E}[X_iX_j] + \sum_{i=1}^n Q_{ii}\mathbb{E}[X_i^2]\\
&= \sum_{i,j\in[n]:i\neq j} Q_{ij}\frac{m_i}{m}\frac{m_j}{m}\left(1 - \frac{m-r}{r(m-1)}\right)\\
&+ \sum_{i=1}^n Q_{ii}\left(\frac{m_i}{m}\right)^2\left(1 - \frac{m-r}{r(m-1)} + \frac{m(m-r)}{rm_i(m-1)}\right)\\
&= \sum_{i,j\in[n]} Q_{ij}x_i^*x_j^*\left(1 - \frac{m-r}{r(m-1)}\right) + \frac{m-r}{r(m-1)}\sum_{i=1}^n Q_{ii}x_i^*\\
&\leq f(x^*) - \frac{m-r}{r(m-1)}f_{\min,\Delta_n} + \frac{m-r}{r(m-1)}\max_{i\in[n]} Q_{ii}\\
&\leq f(x^*) - \frac{m-r}{r(m-1)}f_{\min,\Delta_n} + \frac{m-r}{r(m-1)}f_{\max,\Delta_n}.
\end{aligned}$$

Hence, we obtain

$$\mathbb{E}\left[f(X)\right] - f_{\Delta(n,m)} = \mathbb{E}\left[f(X)\right] - f(x^*) \leq \frac{m-r}{r(m-1)}(f_{\max,\Delta_n} - f_{\min,\Delta_n}).$$

Using Lemma 3.1, we can conclude the proof. □

For any quadratic $f$ with integer coefficients, Vavasis [95] shows that there always exists a rational global minimizer $x^*$ for problem (1.21); see Appendix C.1 for details.

**Example 3.6.** *There exist polynomials of degree larger than* 2 *for which problem (1.21) does not have any rational global minimizer. This is the case, for instance, for the polynomial* $f(x) = 2x_1^3 - x_1 \left(\sum_{i=1}^n x_i\right)^2$*, whose global minimizer always has the irrational component* $x_1 = 1/\sqrt{6}$*.*

We now assume the rational minimizer $x^*$ for problem (1.21) has denominator $m$. Then our next result gives an upper bound for the error estimate $f_{\Delta(n,r)} - f_{\min,\Delta_n}$, in terms of this denominator $m$.

**Theorem 3.7.** *Let* $f = x^T Q x \in \mathcal{H}_{n,2}$*, and let* $x^*$ *be a global minimizer of* $f$ *over* $\Delta_n$ *with denominator* $m$*. For any integer* $r \geq 1$*, one has*

$$f_{\Delta(n,r)} - f_{\min,\Delta_n} \leq \frac{m}{r^2}\left(f_{\max,\Delta_n} - f_{\min,\Delta_n}\right).$$

Before proceeding with the proof, we note that one may give an upper bound on $m$ in terms of $Q$, when $Q$ has integer entries. However, the best-known upper bounds on $m$ are exponential in $n$ in general. This means that Theorem 3.7 does not yield a PTAS (recall Definition 2.2) for quadratic polynomial optimization over the standard simplex, but our interest is on the dependence of the error bound on the parameter $r$, when the polynomial $f$ is fixed. For the upper bound on $m$, see Appendix C.2 for details.

The proof of Theorem 3.7 uses the following easy fact (whose proof is omitted).

**Lemma 3.8.** *Let* $r, k, m \geq 1$ *be integers such that* $(k-1)m < r \leq km$*. Then,*

$$\frac{km - r}{km - 1} \leq \frac{m}{r}.$$

**Proof.** *(Proof of Theorem 3.7)* Let $k \geq 1$ be an integer such that

$$(k-1)m < r \leq km.$$

We apply Theorem 3.5 to $r$ and $km$ (instead of $m$) and obtain that

$$f_{\Delta(n,r)} - f_{\Delta(n,km)} \leq \frac{km - r}{r(km-1)}(f_{\max,\Delta_n} - f_{\min,\Delta_n}).$$

Now, observe that $f_{\Delta(n,km)} = f_{\Delta(n,m)} = f_{\min,\Delta_n}$, since $x^* \in \Delta(n,m) \subseteq \Delta(n,km) \subseteq \Delta_n$, and use the inequality from Lemma 3.8. □

As a direct application of Theorem 3.7, we see that the convergence ratio $\rho_r(f)$ in (1.17) is in the order $O(1/r^2)$, where the constant depends only on the denominator of a rational global minimizer.

**Corollary 3.9.** *For any quadratic polynomial $f \in \mathcal{H}_{n,2}$, $\rho_r(f) = O(1/r^2)$.*

Moreover, the results of Theorems 3.5 and 3.7 refine the known error estimate from Theorem 2.1 (i), which shows that $\rho_r(f) \leq \frac{1}{r}$. To see it, use Theorem 3.5 and the fact that $\frac{m-r}{r(m-1)} \leq \frac{1}{r}$ if $1 \leq r \leq m$, and use Theorem 3.7 and the inequality $\frac{m}{r^2} \leq \frac{1}{r}$ in the case $r \geq m$.

The following example shows that the inequality in Theorem 3.5 can be tight.

**Example 3.10.** *[Example 2.10 continued] Consider the quadratic polynomial $f = \sum_{i=1}^n x_i^2$. Recall that $f_{\max,\Delta_n} = 1$ and $f_{\min,\Delta_n} = \frac{1}{n}$, which is attained at $x = \frac{1}{n}\mathbf{e}$. Moreover, for any integer $r \leq n$, we have $f_{\Delta(n,r)} = \frac{1}{r}$. Thus, we have*

$$f_{\Delta(n,r)} - f_{\min,\Delta_n} = \frac{n-r}{r(n-1)}(f_{\max,\Delta_n} - f_{\min,\Delta_n}) = \frac{m-r}{r(m-1)}(f_{\max,\Delta_n} - f_{\min,\Delta_n}).$$

*Hence, for this example, the result in Theorem 3.5 is tight, while the result in Theorem 2.1 (i) is not tight.*

### 3.3.2 The cubic and square-free cases

For the minimization of cubic and square-free polynomials over the standard simplex, we can show the following analogue of Theorem 3.5.

**Theorem 3.11.** *(i) Let $f \in \mathcal{H}_{n,3}$. Given integers $r, m$ satisfying $1 \leq r \leq m$ and $m \geq 3$, one has*

$$f_{\Delta(n,r)} - f_{\Delta(n,m)} \leq \frac{(m-r)(4mr-2m-2r)}{r^2(m-1)(m-2)}\left(f_{\max,\Delta_n} - f_{\min,\Delta_n}\right).$$

*(ii) Let $f \in \mathcal{H}_{n,d}$ be a square-free polynomial. Given integers $r, m$ satisfying*

$$1 \leq r \leq m \text{ and } m \geq d,$$

*one has*

$$f_{\Delta(n,r)} - f_{\Delta(n,m)} \leq \left(1 - \frac{r^{\underline{d}}}{r^d}\frac{m^d}{m^{\underline{d}}}\right)\left(f_{\max,\Delta_n} - f_{\min,\Delta_n}\right).$$

**Proof.** As in the proof of Theorem 3.5, let $x^* \in \Delta(n,m)$ be a minimizer of $f$ over $\Delta(n,m)$, i.e., $f(x^*) = f_{\Delta(n,m)}$, and set $m_i = mx_i^*$ for $i \in [n]$. Consider the random variables $X_i$ defined in (3.1) and (3.2), so that $X = (X_1, X_2, \ldots, X_n)$ takes its values in $\Delta(n,r)$.

(i) First we consider the case when $f$ is a homogeneous polynomial of degree 3. Write $f$ as

$$f = \sum_{i=1}^{n} f_i x_i^3 + \sum_{1 \le i < j \le n} (f_{ij} x_i x_j^2 + g_{ij} x_i^2 x_j) + \sum_{1 \le i < j < k \le n} f_{ijk} x_i x_j x_k.$$

By Corollary 3.4, for any $i,j,k \in [n]$, one has

$$\begin{aligned}
&\mathbb{E}[X_i^3] = \left(\frac{m_i}{m}\right)^3 \left[1 - (m-r)\frac{3mr - 2(m+r)}{r^2(m-1)(m-2)} + \cdots \right.\\
&\left. +(m-r)\frac{3rm_i m^2 - 3m_i m^2 + m^3 - 2rm^2}{r^2 m_i^2 (m-1)(m-2)}\right]\\
&\mathbb{E}[X_i^2 X_j] = \left(\frac{m_i}{m}\right)^2 \frac{m_j}{m}\left[1 - (m-r)\frac{3mr - 2(m+r)}{r^2(m-1)(m-2)} + \cdots \right.\\
&\left. +(m-r)\frac{(r-1)m^2}{r^2 m_i (m-1)(m-2)}\right]\\
&\mathbb{E}[X_i X_j X_k] = \frac{m_i}{m}\frac{m_j}{m}\frac{m_k}{m}\left[1 - (m-r)\frac{3mr - 2(m+r)}{r^2(m-1)(m-2)}\right].
\end{aligned}$$

Therefore, one obtains

$$\begin{aligned}
&\mathbb{E}\left[f(X)\right]\\
= \ & \sum_i f_i \mathbb{E}[X_i^3] + \sum_{i<j}\left(f_{ij}\mathbb{E}[X_i X_j^2] + g_{ij}\mathbb{E}[X_i^2 X_j]\right) + \sum_{i<j<k} f_{ijk}\mathbb{E}[X_i X_j X_k]\\
= \ & f(x^*)\left[1 - (m-r)\frac{3mr - 2(m+r)}{r^2(m-1)(m-2)}\right] + \frac{m-r}{r^2(m-1)(m-2)}\sigma, \qquad (3.3)
\end{aligned}$$

where we set

$$\sigma = \sum_{i=1}^{n} f_i \frac{m_i}{m}(3m_i r - 3m_i + m - 2r) + m(r-1)\sum_{i<j}(f_{ij} + g_{ij})\frac{m_i}{m}\frac{m_j}{m}. \qquad (3.4)$$

As in [21], by evaluating $f$ at $\mathbf{e}_i$ and $(\mathbf{e}_i + \mathbf{e}_j)/2$, we obtain respectively the relations:

$$f_{\min,\Delta_n} \le f_i \le f_{\max,\Delta_n}, \qquad (3.5)$$

$$f_i + f_j + f_{ij} + g_{ij} \le 8 f_{\max,\Delta_n}. \qquad (3.6)$$

Using (3.6), we obtain

$$\begin{aligned}\sum_{i<j}(f_{ij}+g_{ij})\frac{m_i}{m}\frac{m_j}{m} &\leq \sum_{i<j}(8f_{\max,\Delta_n}-f_i-f_j)\frac{m_i}{m}\frac{m_j}{m}\\ &= 8f_{\max,\Delta_n}\sum_{i<j}\frac{m_i}{m}\frac{m_j}{m}-\sum_{i=1}^n f_i\frac{m_i}{m}\left(1-\frac{m_i}{m}\right).\end{aligned}$$

We use this inequality together with (3.5) to upper bound the term $\sigma$ from (3.4):

$$\begin{aligned}\sigma &\leq \sum_{i=1}^n f_i\frac{m_i}{m}(4m_ir-4m_i+2m-2r-mr)+8m(r-1)f_{\max,\Delta_n}\sum_{i<j}\frac{m_i}{m}\frac{m_j}{m}\\ &= 4m(r-1)\left(\sum_{i=1}^n f_i\left(\frac{m_i}{m}\right)^2+2f_{\max,\Delta_n}\sum_{i<j}\frac{m_i}{m}\frac{m_j}{m}\right)+(2m-2r-mr)\sum_{i=1}^n f_i\frac{m_i}{m}\\ &\leq 4m(r-1)f_{\max,\Delta_n}\left(\sum_{i=1}^n\left(\frac{m_i}{m}\right)^2+2\sum_{i<j}\frac{m_i}{m}\frac{m_j}{m}\right)+2(m-r)\sum_{i=1}^n f_i\frac{m_i}{m}\\ &\quad -mr\sum_{i=1}^n f_i\frac{m_i}{m}\\ &\leq 4m(r-1)f_{\max,\Delta_n}+2(m-r)f_{\max,\Delta_n}-mrf_{\min,\Delta_n}\\ &= (4mr-2m-2r)f_{\max,\Delta_n}-mrf_{\min,\Delta_n}.\end{aligned}$$

We can now upper bound the quantity $\mathbb{E}[f(X)]$ from (3.3) as follows:

$$\mathbb{E}[f(X)]\leq f(x^*)+\frac{(m-r)(4mr-2m-2r)}{r^2(m-1)(m-2)}(f_{\max,\Delta_n}-f_{\min,\Delta_n}).$$

Together with Lemma 3.1, this now concludes the proof of Theorem 3.11 (i).

(ii) We now consider the case when $f$ is a homogeneous square-free polynomial of degree $d$. Say, $f=\sum_{I\subseteq[n],|I|=d}f_Ix^I$. By Corollary 3.4, one has

$$\mathbb{E}[X^I] = \frac{r^{\underline{d}}}{r^d}\frac{\prod_{i\in I}m_i}{m^{\underline{d}}}=\frac{r^{\underline{d}}}{r^d}\frac{m^d}{m^{\underline{d}}}\prod_{i\in I}\frac{m_i}{m}$$

and thus

$$\mathbb{E}[f(X)]=\sum_{I\subseteq[n],|I|=d}f_I\mathbb{E}[X^I]=\frac{r^{\underline{d}}}{r^d}\frac{m^d}{m^{\underline{d}}}f(x^*)=\frac{r^{\underline{d}}}{r^d}\frac{m^d}{m^{\underline{d}}}f_{\Delta(n,m)}.$$

Therefore,

$$\begin{aligned}\mathbb{E}[f(X)] - f_{\Delta(n,m)} &= -\left(1 - \frac{r^{\underline{d}}}{r^d}\frac{m^d}{m^{\underline{d}}}\right) f_{\Delta(n,m)} \le -\left(1 - \frac{r^{\underline{d}}}{r^d}\frac{m^d}{m^{\underline{d}}}\right) f_{\min,\Delta_n} \\ &\le \left(1 - \frac{r^{\underline{d}}}{r^d}\frac{m^d}{m^{\underline{d}}}\right)(f_{\max,\Delta_n} - f_{\min,\Delta_n}).\end{aligned}$$

Here, for the last inequality we have used the fact that $f_{\max,\Delta_n} \ge 0$ (since $f(\mathbf{e}_i) = 0$ for any $i \in [n]$). Together with Lemma 3.1, this concludes the proof of Theorem 3.11 (ii). $\square$

When $f$ is a cubic or square-free polynomial admitting a rational global minimizer in $\Delta_n$, one can show the following result.

**Corollary 3.12.** *(i) Let $f \in \mathcal{H}_{n,3}$ and assume that $f$ has a rational global minimizer with denominator $m \ge 3$ in $\Delta_n$. Then Theorem 3.11 (i) implies Corollary 2.12 for $r \ge 1 + \frac{m-1}{\sqrt{2m-1}}$.*

*(ii) Let $f \in \mathcal{H}_{n,d}$ be a square-free polynomial and assume that $f$ has a rational global minimizer with denominator $m \ge d$ in $\Delta_n$. Then Theorem 3.11 (ii) implies Corollary 2.15*

**Proof.** (i) Now we show how to derive Corollary 2.12 for $r \ge 1 + \frac{m-1}{\sqrt{2m-1}}$ from Theorem 3.11 (i).

When $1 + \frac{m-1}{\sqrt{2m-1}} \le r \le m$, this follows directly from the fact that

$$\frac{(m-r)(4mr-2m-2r)}{r^2(m-1)(m-2)} \le \frac{4}{r} - \frac{4}{r^2}.$$

Assume now $r > m \ge 3$ and $(k-1)m < r \le km$ for some integer $k \ge 2$. It suffices to show the inequality $\frac{(km-r)(4kmr-2km-2r)}{r^2(km-1)(km-2)} \le \frac{4}{r} - \frac{4}{r^2}$ or, equivalently,

$$\varphi(r) := (2km-1)r^2 + (4-6km)r - k^2m^2 + 6km - 4 \ge 0.$$

One can check that the function $\varphi(r)$ is monotonically increasing for $r \ge 1 + \frac{km-1}{2km-1}$ and thus for $r \ge 2$. Hence it suffices to show that $\varphi((k-1)m+1) \ge 0$. If $m \ge 3$ is fixed, then one can check that $\varphi((k-1)m+1)$, as a function of $k$, is monotonically increasing for $k \ge 2$. Therefore, it suffices to show that $\varphi((k-1)m+1) \ge 0$ when $k = 2$ and $m \ge 3$. One can now check that $\varphi((k-1)m+1)$ with $k = 2$, as a function of $m$, is monotonically increasing for $m \ge 3$. Finally, we can conclude that it suffices to show that $\varphi((k-1)m+1) \geq 0$ when $k=2$ and $m=3$, which can be easily checked to hold. Thus we have shown that $\varphi(r) \geq 0$ for any $r > m$.

(ii) To see that Theorem 3.11 (ii) implies Corollary 2.15, consider an integer $k \geq 1$ such that $(k-1)m < r \leq km$ and observe that $1 - \frac{r^{\underline{d}}(km)^d}{r^d(km)^{\underline{d}}} \leq 1 - \frac{r^{\underline{d}}}{r^d}$. ☐

As an application of Theorem 3.11, we can show that $\rho_r(f)$ is in the order $O(1/r^2)$ for cubic polynomials admitting a rational global minimizer over the standard simplex (see Corollary 3.13). The same holds for square-free polynomials as we will see in Section 3.3.3.

**Corollary 3.13.** *Let $f \in \mathcal{H}_{n,3}$ and assume that $f$ has a rational global minimizer with denominator $m \geq 3$ in $\Delta_n$. Then,*

$$\rho_r(f) \leq \frac{12m}{r^2}$$

*and thus $\rho_r(f) = O(1/r^2)$.*

**Proof.** If $1 \leq r \leq m$ then, by using Theorem 3.11 (i), Lemma 3.8 and the inequality $\frac{4mr-2m-2r}{r(m-2)} \leq \frac{4(m-1)}{m-2}$, we deduce that

$$\rho_r(f) \leq \frac{(m-r)(4mr-2m-2r)}{r^2(m-1)(m-2)} \leq \frac{4m^2}{r^2(m-2)} \leq \frac{12m}{r^2}.$$

Assume now $r > m$ and $(k-1)m < r \leq km$ for some integer $k \geq 2$. Then Theorem 3.11 (i) implies

$$f_{\Delta(n,r)} - f_{\min,\Delta_n} = f_{\Delta(n,r)} - f_{\Delta(n,km)} \leq \frac{(km-r)(4kmr-2km-2r)}{r^2(km-1)(km-2)} \left(f_{\max,\Delta_n} - f_{\min,\Delta_n}\right).$$

One can easily check that $\frac{4kmr-2km-2r}{r(km-2)} \leq 6$ which, together with Lemma 3.8, implies that $\rho_r(f) \leq \frac{6m}{r^2} \leq \frac{12m}{r^2}$. This concludes the proof of Corollary 3.13. ☐

### 3.3.3 The general case

We now study the general fixed-degree polynomial optimization problem over the standard simplex. We first upper bound the range $f_{\Delta(n,r)} - f_{\Delta(n,m)}$ in terms of the range $f_{\max,\Delta_n} - f_{\min,\Delta_n}$.

**Theorem 3.14.** *Let $f \in \mathcal{H}_{n,d}$. For any integers $r, m$ satisfying $1 \le r \le m$ and $m \ge d$, one has*

$$f_{\Delta(n,r)} - f_{\Delta(n,m)} \le \left(1 - \frac{r^{\underline{d}} m^d}{r^d m^{\underline{d}}}\right) \binom{2d-1}{d} d^d (f_{\max,\Delta_n} - f_{\min,\Delta_n}).$$

Note that when $f$ is square-free, we have proved a better bound in Theorem 3.11 (ii).

For the proof of Theorem 3.14, we will use the Vandermonde-Chu identity

$$\left(\sum_{i=1}^n x_i\right)^{\underline{d}} = \sum_{\alpha \in I(n,d)} \frac{d!}{\alpha!} x^{\underline{\alpha}} \quad \forall x \in \mathbb{R}^n \tag{3.7}$$

(see [83] for a proof, or alternatively use induction on $d \ge 1$), as well as the multinomial theorem

$$(\sum_{i=1}^n x_i)^r = \sum_{\alpha \in I(n,r)} \frac{r!}{\alpha!} x^\alpha. \tag{3.8}$$

We will also use the following technical lemma.

**Lemma 3.15.** *Given $\beta \in I(n,d)$, for any integers $r, m$ with $1 \le r \le m$, $m \ge d$ and integers $m_i$ ($i \in [n]$) with $\sum_{i=1}^n m_i = m$, one has*

$$A_\beta := r^{\underline{d}} \left(\prod_{i=1}^n m_i^{\underline{\beta_i}} - \prod_{i=1}^n m_i^{\beta_i}\right) + \sum_{\alpha \in \mathbb{N}^n : \alpha \le \beta, \alpha \ne \beta} \frac{r^{\underline{|\alpha|}} m^{\underline{d}}}{m^{\underline{|\alpha|}}} \prod_{i=1}^n m_i^{\underline{\alpha_i}} S(\beta_i, \alpha_i) \ge 0, \tag{3.9}$$

$$\sum_{\beta \in I(n,d)} \frac{d!}{\beta!} A_\beta = r^d m^{\underline{d}} - r^{\underline{d}} m^d. \tag{3.10}$$

**Proof.** We first prove (3.9). For any $\alpha \in \mathbb{N}^n$ with $|\alpha| \le d$, one can easily check that $\frac{r^{\underline{|\alpha|}}}{r^{\underline{d}}} \ge \frac{m^{\underline{|\alpha|}}}{m^{\underline{d}}}$, that is, $r^{\underline{d}} \le \frac{r^{\underline{|\alpha|}} m^{\underline{d}}}{m^{\underline{|\alpha|}}}$. Hence, one has

$$\begin{aligned} A_\beta &= r^{\underline{d}} \left(\prod_{i=1}^n m_i^{\underline{\beta_i}} - \prod_{i=1}^n m_i^{\beta_i}\right) + \sum_{\alpha \in \mathbb{N}^n : \alpha \le \beta, \alpha \ne \beta} \frac{r^{\underline{|\alpha|}} m^{\underline{d}}}{m^{\underline{|\alpha|}}} \prod_{i=1}^n m_i^{\underline{\alpha_i}} S(\beta_i, \alpha_i) \\ &\ge r^{\underline{d}} \underbrace{\left(\prod_{i=1}^n m_i^{\underline{\beta_i}} - \prod_{i=1}^n m_i^{\beta_i} + \sum_{\alpha \in \mathbb{N}^n : \alpha \le \beta, \alpha \ne \beta} \prod_{i=1}^n m_i^{\underline{\alpha_i}} S(\beta_i, \alpha_i)\right)}_{=: B_\beta} = r^{\underline{d}} B_\beta. \end{aligned}$$

Then we consider the quantity $B_\beta$ and show that $B_\beta = 0$. As $S(\beta_i, \beta_i) = 1$, one can rewrite $B_\beta$ as

$$B_\beta = \sum_{\alpha\in\mathbb{N}^n:\alpha\leq\beta} \prod_{i=1}^{n} m_i^{\underline{\alpha_i}} S(\beta_i, \alpha_i) - \prod_{i=1}^{n} m_i^{\beta_i}.$$

Applying Lemma A.3 (with $(m_i, \beta_i)$ in place of $(r, d)$), we have

$$m_i^{\beta_i} = \sum_{\alpha_i=0}^{\beta_i} m_i^{\underline{\alpha_i}} S(\beta_i, \alpha_i), \quad \text{implying} \quad \prod_{i=1}^{n} m_i^{\beta_i} = \sum_{\alpha\in\mathbb{N}^n:\alpha\leq\beta} \prod_{i=1}^{n} m_i^{\underline{\alpha_i}} S(\beta_i, \alpha_i),$$

which shows that $B_\beta = 0$, and thus $A_\beta \geq 0$, which concludes the proof of (3.9).

We now show (3.10). By the definition (3.9), one has

$$\begin{aligned}
\sum_{\beta\in I(n,d)} \frac{d!}{\beta!} A_\beta \quad &= \quad \underbrace{\sum_{\beta\in I(n,d)} \frac{d!}{\beta!} r^{\underline{d}} \left( \prod_{i=1}^{n} m_i^{\underline{\beta_i}} - \prod_{i=1}^{n} m_i^{\beta_i} \right)}_{=:C_1} \\
&\quad + \quad \underbrace{\sum_{\beta\in I(n,d)} \frac{d!}{\beta!} \sum_{\alpha\in\mathbb{N}^n:\alpha\leq\beta,\alpha\neq\beta} \frac{r^{\underline{|\alpha|}} m^{\underline{d}}}{m^{\underline{|\alpha|}}} \prod_{i=1}^{n} m_i^{\underline{\alpha_i}} S(\beta_i, \alpha_i)}_{=:C_2}.
\end{aligned}$$

On the one hand, using the Vandermonde-Chu identity (3.7), the multinomial theorem (3.8) and the identity $\sum_{i=1}^{n} m_i = m$, we find

$$C_1 = r^{\underline{d}}(m^{\underline{d}} - m^d).$$

On the other hand, we may exchange the summations in the definition of $C_2$, by recalling that $S(\beta_i, \alpha_i) = 0$ if $\alpha_i > \beta_i$, and noting that $\alpha \leq \beta$, $\alpha \neq \beta$, and $\beta \in I(n, d)$ implies that $\alpha \in I(n, k)$ for some $k < d$. This allows us to remove the conditions

$\alpha \leq \beta$ and $\alpha \neq \beta$ in the summation, and we obtain:

$$\begin{aligned} C_2 &= m^{\underline{d}} \sum_{k=1}^{d-1} \sum_{\alpha \in I(n,k)} \frac{r^{\underline{|\alpha|}}}{m^{\underline{|\alpha|}}} \prod_{i=1}^{n} m_i^{\underline{\alpha_i}} \left( \sum_{\beta \in I(n,d)} \frac{d!}{\beta!} \prod_{i=1}^{n} S(\beta_i, \alpha_i) \right) \\ &= m^{\underline{d}} \sum_{k=1}^{d-1} \frac{r^{\underline{k}}}{m^{\underline{k}}} S(d,k) \left( \sum_{\alpha \in I(n,k)} \frac{k!}{\alpha!} \prod_{i=1}^{n} m_i^{\underline{\alpha_i}} \right) && \text{[using Lemma A.2]} \\ &= m^{\underline{d}} \sum_{k=1}^{d-1} r^{\underline{k}} S(d,k) && \text{[using Vandermonde-Chu identity (3.7)]} \\ &= m^{\underline{d}} (r^d - r^{\underline{d}}) && \text{[using Lemma A.3]} \end{aligned}$$

We can now conclude that $\sum_{\beta \in I(n,d)} \frac{d!}{\beta!} A_\beta = C_1 + C_2 = r^d m^{\underline{d}} - r^{\underline{d}} m^d$. $\square$

Now we are ready to prove Theorem 3.14.

**Proof.** *(of Theorem 3.14)* Let $x^* \in \Delta(n,m)$ be a minimizer of $f$ over $\Delta(n,m)$, i.e., $f(x^*) = f_{\Delta(n,m)}$. Set $m_i = m x_i^*$ for $i \in [n]$. Let the random variables $X_i$ be defined as in (3.1) and (3.2), so that the random variable $X = (X_1, X_2, \ldots, X_n)$ takes its values in $\Delta(n,r)$. By Corollary 3.4 we have for $\beta \in I(n,d)$:

$$\mathbb{E}[X^\beta] = \frac{1}{r^d} \sum_{\alpha \in \mathbb{N}^n : \alpha \leq \beta} \frac{r^{\underline{|\alpha|}}}{m^{\underline{|\alpha|}}} \prod_{i=1}^{n} m_i^{\underline{\alpha_i}} S(\beta_i, \alpha_i).$$

Then, as $S(\beta_i, \beta_i) = 1$, we can rewrite

$$\begin{aligned} \mathbb{E}[X^\beta] &= \frac{1}{r^d} \frac{r^{\underline{d}}}{m^{\underline{d}}} \prod_{i=1}^{n} m_i^{\underline{\beta_i}} + \underbrace{\frac{1}{r^d} \sum_{\alpha \in \mathbb{N}^n : \alpha \leq \beta, \alpha \neq \beta} \frac{r^{\underline{|\alpha|}}}{m^{\underline{|\alpha|}}} \prod_{i=1}^{n} m_i^{\underline{\alpha_i}} S(\beta_i, \alpha_i)}_{=: D_\beta} \\ &= \prod_{i=1}^{n} \left( \frac{m_i}{m} \right)^{\beta_i} \left[ \frac{r^{\underline{d}}}{r^d} \frac{m^d}{m^{\underline{d}}} + \frac{r^{\underline{d}}}{r^d} \frac{m^d}{m^{\underline{d}}} \left( \prod_{i=1}^{n} \frac{m_i^{\underline{\beta_i}}}{m_i^{\beta_i}} - 1 \right) \right] + D_\beta \\ &= \underbrace{\prod_{i=1}^{n} \left( \frac{m_i}{m} \right)^{\beta_i} \frac{r^{\underline{d}}}{r^d} \frac{m^d}{m^{\underline{d}}}}_{=: T_1} + \underbrace{\frac{r^{\underline{d}}}{r^d m^{\underline{d}}} \left( \prod_{i=1}^{n} m_i^{\underline{\beta_i}} - \prod_{i=1}^{n} m_i^{\beta_i} \right) + D_\beta}_{=: T_2} \\ &= T_1 + T_2 = (x^*)^\beta \frac{r^{\underline{d}} m^d}{r^d m^{\underline{d}}} + \frac{A_\beta}{r^d m^{\underline{d}}}. \end{aligned}$$

For the above last equality, note that, since $x_i^* = \frac{m_i}{m}$, one has

$$T_1 = (x^*)^\beta \frac{r^{\underline{d}} m^d}{r^d m^{\underline{d}}}$$

and using the definition of $A_\beta$ in (3.9), we can write

$$T_2 = \frac{A_\beta}{r^d m^{\underline{d}}}.$$

Thus we obtain

$$\begin{aligned}
\mathbb{E}[f(X)] &= \mathbb{E}[\sum_{\beta \in I(n,d)} f_\beta X^\beta] = \sum_{\beta \in I(n,d)} f_\beta \mathbb{E}[X^\beta] \\
&= \frac{r^{\underline{d}}}{r^d} \frac{m^d}{m^{\underline{d}}} f(x^*) + \frac{1}{r^d m^{\underline{d}}} \sum_{\beta \in I(n,d)} f_\beta A_\beta.
\end{aligned}$$

Therefore, we have

$$r^d m^{\underline{d}} (\mathbb{E}[f(X)] - f(x^*)) = (r^{\underline{d}} m^d - r^d m^{\underline{d}}) f(x^*) + \sum_{\beta \in I(n,d)} f_\beta A_\beta.$$

We now upper bound the two terms $(r^{\underline{d}} m^d - r^d m^{\underline{d}}) f(x^*)$ and $\sum_{\beta \in I(n,d)} f_\beta A_\beta$. First, since $r^{\underline{d}} m^d - r^d m^{\underline{d}} < 0$ and $f(x^*) \geq \min_{\beta \in I(n,d)} f_\beta$ (see (2.10)), one obtains

$$(r^{\underline{d}} m^d - r^d m^{\underline{d}}) f(x^*) \leq (r^{\underline{d}} m^d - r^d m^{\underline{d}}) \left( \min_{\beta \in I(n,d)} f_\beta \frac{\beta!}{d!} \right). \tag{3.11}$$

Second, using the fact that $A_\beta \geq 0$ (by Lemma 3.15), one obtains

$$\sum_{\beta \in I(n,d)} f_\beta A_\beta \leq \left( \max_{\beta \in I(n,d)} f_\beta \frac{\beta!}{d!} \right) \sum_{\beta \in I(n,d)} \frac{d!}{\beta!} A_\beta.$$

Using the identity $\sum_{\beta \in I(n,d)} \frac{d!}{\beta!} A_\beta = r^d m^{\underline{d}} - r^{\underline{d}} m^d$ (see (3.10)), one can obtain

$$\sum_{\beta \in I(n,d)} f_\beta A_\beta \leq \left( \max_{\beta \in I(n,d)} f_\beta \frac{\beta!}{d!} \right) (r^d m^{\underline{d}} - r^{\underline{d}} m^d).$$

Combining with (3.11), this implies

$$r^d m^{\underline{d}}(\mathbb{E}[f(X)] - f(x^*)) \leq (r^d m^{\underline{d}} - r^{\underline{d}} m^d) \left( \max_{\beta \in I(n,d)} f_\beta \frac{\beta!}{d!} - \min_{\beta \in I(n,d)} f_\beta \frac{\beta!}{d!} \right).$$

Using Theorem 2.6, Lemma 3.1 and the fact that $f(x^*) = f_{\Delta(n,m)}$, we finally obtain

$$\begin{aligned} r^d m^{\underline{d}}(f_{\Delta(n,r)} - f_{\Delta(n,m)}) &\leq r^d m^{\underline{d}}(\mathbb{E}[f(X)] - f(x^*)) \\ &\leq (r^d m^{\underline{d}} - r^{\underline{d}} m^d) \binom{2d-1}{d} d^d (f_{\max,\Delta_n} - f_{\min,\Delta_n}), \end{aligned}$$

which concludes the proof of Theorem 3.14. □

In what follows we assume that $f \in \mathcal{H}_{n,d}$ has a rational global minimizer $x^*$ with denominator $m$, i.e., $x^* \in \Delta(n,m)$, so that $f_{\min,\Delta_n} = f_{\Delta(n,m)}$.

First, observe that Theorem 3.14 refines the result from Theorem 2.1 (ii) (which follows from the fact that $1 - \frac{r^{\underline{d}}(km)^d}{r^d(km)^{\underline{d}}} \leq 1 - \frac{r^{\underline{d}}}{r^d}$ for any $k \geq 1$).

Next, we show as an application of Theorem 3.14 that the ratio $\rho_r(f)$ is in the order $O(1/r^2)$.

**Corollary 3.16.** *Let $f \in \mathcal{H}_{n,d}$ and assume that there exists a rational global minimizer for problem (1.21). Then, $\rho_r(f) = O(1/r^2)$.*

For the proof of Corollary 3.16, we need the following notation. Consider the univariate polynomial $(x-1)(x-2)\cdots(x-d+1)$ (in the variable $x$), which can be written as

$$\begin{aligned} &(x-1)(x-2)\cdots(x-d+1) \\ =\ & x^{d-1} - a_{d-2}x^{d-2} + a_{d-3}x^{d-3} + \cdots + (-1)^{d-1}a_0 \\ =\ & x^{d-1} + p(x), \end{aligned} \tag{3.12}$$

setting

$$p(x) = \sum_{i=0}^{d-2} (-1)^{d-1-i} a_i x^i, \tag{3.13}$$

where $a_i$ are positive integers depending only on $d$ for any $i \in \{0, 1, \ldots, d-2\}$. We also need the following lemma.

**Lemma 3.17.** *Let $r, m$ and $k$ be integers satisfying*

$$m \geq d,\ \ k \geq 1 \text{ and } (k-1)m < r \leq km.$$

*Then one has*

$$1 - \frac{r^{\underline{d}}(km)^d}{r^d(km)^{\underline{d}}} \leq \frac{m}{r^2}c_d,$$

*for some constant $c_d$ depending only on $d$.*

**Proof.** Based on (3.13), one can write

$$1 - \frac{r^{\underline{d}}(km)^d}{r^d(km)^{\underline{d}}} = \underbrace{\frac{(km)^{d-1}}{(km-1)(km-2)\cdots(km-d+1)}}_{=:\sigma_0(r,km)} \underbrace{\left[\frac{p(km)}{(km)^{d-1}} - \frac{p(r)}{r^{d-1}}\right]}_{=:\sigma_1(r,km)}.$$

First we consider the term $\sigma_0(r, km)$. For any integer $i \in \{1, \ldots, d-1\}$, as $k \geq 1$ and $m \geq d$, we have that $km(d-1) \geq id$, which implies $\frac{km}{km-i} \leq d$. Hence, one has $\sigma_0(r, km) \leq d^{d-1}$.

Next we consider the term $\sigma_1(r, km)$. Recalling (3.12), we can write $\sigma_1(r, km)$ as $\sigma_1(r, km) = \sum_{i=0}^{d-2}(-1)^{d-1-i}a_i\left(\frac{1}{(km)^{d-1-i}} - \frac{1}{r^{d-1-i}}\right)$. Since $r \leq km$, then $\frac{1}{km} \leq \frac{1}{r}$ and $\frac{1}{(km)^{d-1-i}} \leq \frac{1}{r^{d-1-i}}$ for any $i \in \{0, 1, \ldots, d-2\}$. This gives:

$$\sigma_1(r, km) \leq \sum_{i=0}^{d-2} a_i\left(\frac{1}{r^{d-1-i}} - \frac{1}{(km)^{d-1-i}}\right). \tag{3.14}$$

Then, we consider the term $\frac{1}{r^{d-1-i}} - \frac{1}{(km)^{d-1-i}}$ (for any $i \in \{0, 1, \ldots, d-2\}$) in (3.14). For any integer $s \in [d-1]$, we have

$$\frac{1}{r^s} - \frac{1}{(km)^s} = \frac{(km)^s - r^s}{r^s(km)^s} = D_1 \cdot D_2,$$

setting

$$\begin{aligned} D_1 &= \frac{km - r}{kmr}, \\ D_2 &= \frac{(km)^{s-1} + (km)^{s-2}r + \cdots + r^{s-1}}{r^{s-1}(km)^{s-1}}. \end{aligned}$$

On the one hand, one has $D_1 \leq \frac{km-r}{r(km-1)} \leq \frac{m}{r^2}$, where the second inequality follows by Lemma 3.8. On the other hand, observe that for any $i, j \in \{0, 1, \ldots, s-1\}$ with $i + j = s - 1$, one has $(km)^i r^j \leq (km)^{s-1} r^{s-1}$. Hence, $D_2 \leq s \leq d - 1$. That is,

$$\frac{1}{r^s} - \frac{1}{(km)^s} \leq \frac{m(d-1)}{r^2}.$$

Using this in (3.14), we find that $\sigma_1(r, km) \leq \frac{m(d-1)}{r^2} \sum_{i=0}^{d-2} a_i$. From (3.12) and (3.13), we know that the term $(d-1)(\sum_{i=0}^{d-2} a_i)$ is a constant $c_d$ that depends only on $d$. This concludes the proof. □

We can now prove Corollary 3.16.

**Proof.** *(proof of Corollary 3.16)* Let $x^* \in \Delta(n, m)$ be a rational global minimizer of $f$ over $\Delta_n$. Let $r \geq d$ and let $k \geq 1$ be an integer such that $(k-1)m < r \leq km$. Using Theorem 3.14 (applied to $r$ and $km$ (instead of $m$)), we obtain that

$$f_{\Delta(n,r)} - f_{\min,\Delta_n} = f_{\Delta(n,r)} - f_{\Delta(n,km)} \leq \left(1 - \frac{r^{\underline{d}}(km)^d}{r^d (km)^{\underline{d}}}\right) \binom{2d-1}{d} d^d (f_{\max,\Delta_n} - f_{\min,\Delta_n}).$$

Combining with Lemma 3.17, one can conclude. □

## 3.4 Concluding remarks

As explained in Section 3.1, the analysis presented in this chapter is essentially a modification of the analysis in Chapter 2, in the sense that one discrete distribution on $\Delta(n, r)$ is replaced by another one. However, the technical details turn out to be a bit more involved.

Having said that the result in Theorem 3.7 does not imply a PTAS for quadratic $f$, since the estimates on $m$ are exponential in $n$ in general. Moreover, note that for non-quadratic $f$ the analysis in this chapter does not imply a PTAS either, due to the restrictive assumption of existence of a rational global minimizer. It is not clear at this time if this assumption is an artefact of our analysis using the hypergeometric distribution, or if there exist instances of problem (1.21) where all global minimizers are irrational and $\rho_r(f) = \Omega(1/r)$.

Thus, it remains an open problem to determine whether $\rho_r(f) = O(1/r^2)$, which can be an interesting question for future research.

# Chapter 4

# The hierarchy of lower bounds based on Pólya's theorem

## 4.1 Introduction

Let $f$ be a homogeneous polynomial of degree $d$. As we saw earlier in Theorem 1.1, Pólya [82] proves that if $f$ is positive on $\Delta_n$, then $(\sum_{i=1}^n x_i)^r f(x)$ has nonnegative coefficients for $r$ large enough. As mentioned in Section 1.2.1, based on Pólya's representation theorem, one can define the parameter $f_{\min}^{(r-d)}$ for any integer $r \geq d$ as

$$f_{\min}^{(r-d)} = \max \lambda \ \text{ s.t. } \left(\sum_{i=1}^n x_i\right)^{r-d} \left(f - \lambda \left(\sum_{i=1}^n x_i\right)^d\right) \in \mathbb{R}_+[x]$$

(recall (1.8)). Since $f_{\min,\Delta_n}$ can be equivalently formulated as

$$f_{\min,\Delta_n} = \sup \quad \lambda \ \text{ s.t. } \ f(x) - \lambda \left(\sum_{i=1}^n x_i\right)^d \geq 0 \ \ \forall x \in \mathbb{R}_+^n,$$

the parameters $f_{\min}^{(r-d)}$ for $r \geq d$ form a hierarchy of lower bounds for $f_{\min,\Delta_n}$, i.e.,

$$f_{\min}^{(0)} \leq f_{\min}^{(1)} \leq \cdots \leq f_{\min}^{(r-d)} \leq \cdots \leq f_{\min,\Delta_n}.$$

Recall that

$$f_{\Delta(n,r)} = \min \ f(x) \text{ s.t. } x \in \Delta(n,r) = \{x \in \Delta_n : rx \in \mathbb{N}^n\}$$

is an upper bound for $f_{\min,\Delta_n}$. Then, one has

$$f_{\min}^{(0)} \leq f_{\min}^{(1)} \leq \cdots \leq f_{\min}^{(r-d)} \leq \cdots \leq f_{\min,\Delta_n} \leq f_{\Delta(n,r)} \leq f_{\max,\Delta_n}.$$

The idea of applying Pólya's representation theorem to construct hierarchical approximations is first introduced by Parrilo [79, 80] in copositive optimization, and it has been widely used since then, see, e.g., [14, 20, 21, 58] and the references therein.

In this chapter, we consider $f_{\min}^{(r-d)}$ together with $f_{\Delta(n,r)}$, and focus on their difference $f_{\Delta(n,r)} - f_{\min}^{(r-d)}$. To be more precise, we study upper bounds for $f_{\Delta(n,r)} - f_{\min}^{(r-d)}$ in terms of the range $f_{\max,\Delta_n} - f_{\min,\Delta_n}$. In fact, upper bounds for $f_{\Delta(n,r)} - f_{\min,\Delta_n}$ and $f_{\min,\Delta_n} - f_{\min}^{(r-d)}$ have been shown by De Klerk et al. [21], and by adding them up one can easily derive upper bounds for $f_{\Delta(n,r)} - f_{\min}^{(r-d)}$. However, by considering directly the range $f_{\Delta(n,r)} - f_{\min}^{(r-d)}$ we can show sharper results. We consider the cases when $f$ is quadratic in Theorem 4.2, when $f$ is cubic or square-free in Theorem 4.5, and when $f$ is general fixed-degree in Theorem 4.9, and our results refine the respective results from Theorems 4.3, 4.6, and 4.10.

Our main tool is the explicit formula of $f_{\min}^{(r-d)}$ as given in Lemma 4.1 below. We now start by considering how to compute $f_{\min}^{(r-d)}$.

Note that for fixed $r \geq d$, $f_{\min}^{(r-d)}$ can be computed via a linear program in the variable $\lambda$, obtained by checking the nonnegativity for the coefficients of the monomials $x^\alpha$ in the polynomial

$$\left(\sum_{i=1}^n x_i\right)^{r-d} \left(f - \lambda \left(\sum_{i=1}^n x_i\right)^d\right).$$

Based on this, we give an explicit formula for $f_{\min}^{(r-d)}$, which also follows from [83, relation (3)]; note that the quadratic case of this formula has also been observed in [81, 86, 100].

**Lemma 4.1.** *[83] For $f = \sum_{\beta \in I(n,d)} f_\beta x^\beta \in \mathcal{H}_{n,d}$, one has*

$$f_{\min}^{(r-d)} = \min_{\alpha \in I(n,r)} \sum_{\beta \in I(n,d)} f_\beta \frac{\alpha^{\underline{\beta}}}{r^{\underline{d}}}. \tag{4.1}$$

**Proof.** By using the multinomial theorem $(\sum_{i=1}^n x_i)^d = \sum_{\alpha\in I(n,d)} \frac{d!}{\alpha!}x^\alpha$, we obtain

$$\begin{aligned}
&\left(\sum_{i=1}^n x_i\right)^{r-d} f - \lambda \left(\sum_{i=1}^n x_i\right)^r \\
= &\left(\sum_{\gamma\in I(n,r-d)} \frac{(r-d)!}{\gamma!}x^\gamma\right)\left(\sum_{\beta\in I(n,d)} f_\beta x^\beta\right) - \lambda\left(\sum_{\alpha\in I(n,r)} \frac{r!}{\alpha!}x^\alpha\right) \\
= &\sum_{\alpha\in I(n,r)} \left(\sum_{\beta\in I(n,d)} f_\beta \alpha^{\underline{\beta}} \frac{1}{r^{\underline{d}}}\right)\frac{r!}{\alpha!}x^\alpha - \lambda\left(\sum_{\alpha\in I(n,r)} \frac{r!}{\alpha!}x^\alpha\right) \\
= &\sum_{\alpha\in I(n,r)} \left(\sum_{\beta\in I(n,d)} f_\beta \alpha^{\underline{\beta}} \frac{1}{r^{\underline{d}}} - \lambda\right)\frac{r!}{\alpha!}x^\alpha.
\end{aligned}$$

Hence, by definition (1.8), we obtain

$$\begin{aligned}
f_{\min}^{(r-d)} &= \max\ \lambda\ \text{ s.t } \sum_{\beta\in I(n,d)} f_\beta \alpha^{\underline{\beta}} \frac{1}{r^{\underline{d}}} - \lambda \geq 0 \ \ \forall \alpha\in I(n,r) \\
&= \min \sum_{\beta\in I(n,d)} f_\beta \alpha^{\underline{\beta}} \frac{1}{r^{\underline{d}}} \ \text{ s.t } \ \alpha\in I(n,r).
\end{aligned}$$

□

Similarly to the parameter $f_{\Delta(n,r)}$, the computation of $f_{\min}^{(r-d)}$ requires $|I(n,r)| = \binom{n+r-1}{r}$ evaluations of the polynomial $\sum_{\beta\in I(n,d)} f_\beta x^{\underline{\beta}} \frac{1}{r^{\underline{d}}}$ at the points $x \in I(n,r)$, which is polynomial in $n$ for fixed $r$.

## 4.2 Error analysis for this hierarchy

In this section, we show upper bounds for the range $f_{\Delta(n,r)} - f_{\min}^{(r-d)}$. We consider four cases, when $f$ is quadratic, cubic, square-free, and general fixed-degree, respectively, and show that our upper bounds refine the previous known results. Moreover, our results for the quadratic, cubic, and square-free cases refine our result for the general fixed-degree case.

### 4.2.1 The quadratic case

For any quadratic polynomial $f$, we show the following result.

**Theorem 4.2.** *For any quadratic $f = x^T Qx$ and $r \geq 2$, one has*

$$f_{\Delta(n,r)} - f_{\min}^{(r-2)} \leq \frac{1}{r-1}(Q_{\max} - f_{\Delta(n,r)}) \leq \frac{1}{r-1}(f_{\max,\Delta_n} - f_{\min,\Delta_n}), \tag{4.2}$$

*where* $Q_{\max} := \max_{i\in[n]} Q_{ii}$.

**Proof.** By (4.1), we have

$$f_{\min}^{(r-2)} = \min_{\alpha\in I(n,r)} \frac{1}{r(r-1)} \left[ f(\alpha) - \sum_{i=1}^{n} Q_{ii}\alpha_i \right].$$

Hence, $\frac{r-1}{r} f_{\min}^{(r-2)} = \min_{\alpha\in I(n,r)} \left[ f(\frac{\alpha}{r}) - \sum_{i=1}^{n} Q_{ii} \frac{\alpha_i}{r}\frac{1}{r} \right]$. We obtain

$$\frac{r-1}{r} f_{\min}^{(r-2)} \geq \min_{\alpha\in I(n,r)} f(\frac{\alpha}{r}) - \max_{\alpha\in I(n,r)} \frac{1}{r} \sum_{i=1}^{n} Q_{ii}\frac{\alpha_i}{r} = f_{\Delta(n,r)} - \frac{1}{r} Q_{\max}. \tag{4.3}$$

One can easily obtain the first inequality in (4.2) using (4.3). For the second inequality in (4.2), we use the fact that $Q_{\max} \leq f_{\max,\Delta_n}$ (since $Q_{ii} = f(\mathbf{e}_i) \leq f_{\max,\Delta_n}$ for $i \in [n]$), as well as the fact that $f_{\Delta(n,r)} \geq f_{\min,\Delta_n}$. □

Now we point out that our result (4.2) refines the relevant result of [21]. De Klerk et al. [21] show the following theorem.

**Theorem 4.3.** *[21, Theorem 3.2] Suppose $f \in \mathcal{H}_{n,2}$ and $r \geq 2$. Then*

$$f_{\min,\Delta_n} - f_{\min}^{(r-2)} \leq \frac{1}{r-1}(f_{\max,\Delta_n} - f_{\min,\Delta_n}), \tag{4.4}$$

$$f_{\Delta(n,r)} - f_{\min,\Delta_n} \leq \frac{1}{r}(f_{\max,\Delta_n} - f_{\min,\Delta_n}). \tag{4.5}$$

By adding up (4.4) and (4.5), one gets

$$f_{\Delta(n,r)} - f_{\min}^{(r-2)} \leq \left( \frac{1}{r-1} + \frac{1}{r} \right) (f_{\max,\Delta_n} - f_{\min,\Delta_n}),$$

which is implied by our result (4.2).

Recall in Theorem 3.7, we show that

$$f_{\Delta(n,r)} - f_{\min,\Delta_n} \leq \frac{m}{r^2} \left( f_{\max,\Delta_n} - f_{\min,\Delta_n} \right), \tag{4.6}$$

where $m$ is the denominator of a rational global minimizer of $f$ over $\Delta_n$. Similarly as above, our new upper bound (4.2) implies the upper bound obtained by adding up (4.4) and (4.6).

Moreover, Yildirim [100] considers one hierarchical upper bound of $f_{\min,\Delta_n}$ (when $f$ is quadratic), which is defined by $\min_{k\leq r} f_{\Delta(n,k)}$. One can easily verify that

$$f_{\min}^{(r-2)} \leq f_{\min,\Delta_n} \leq \min_{k\leq r} f_{\Delta(n,k)} \leq f_{\Delta(n,r)}.$$

In [100, Theorem 4.1], Yildirim shows $\min_{k\leq r} f_{\Delta(n,k)} - f_{\min}^{(r-2)} \leq \frac{1}{r-1}(Q_{\max} - f_{\min,\Delta_n})$, which thus also follows our result (4.2).

The following example shows that the upper bound (4.2) can be tight.

**Example 4.4.** *[Example 2.10 continued] Consider the quadratic polynomial $f = \sum_{i=1}^n x_i^2$. Recall that $f_{\min,\Delta_n} = \frac{1}{n}$ (attained at $x = \frac{1}{n}\mathbf{e}$) and $f_{\max,\Delta_n} = 1$ (attained at any standard unit vector). To compute $f_{\Delta(n,r)}$, we write $r$ as $r = kn+s$, where $k \geq 0$ and $0 \leq s < n$, and then*

$$f_{\Delta(n,r)} = \frac{1}{n} + \frac{1}{r^2}\frac{s(n-s)}{n}.$$

*By (4.1), we have*

$$f_{\Delta(n,r)} - f_{\min}^{(r-2)} = \frac{1}{r-1}\left(f_{\max,\Delta_n} - f_{\min,\Delta_n}\right) - \frac{1}{r^2(r-1)}\frac{s(n-s)}{n}.$$

*Hence, for this example, the upper bound (4.2) is tight when $s = 0$.*

### 4.2.2 The cubic case

**Theorem 4.5.** *For any cubic polynomial $f$ and $r \geq 3$, one has*

$$f_{\Delta(n,r)} - f_{\min}^{(r-3)} \leq \frac{4r}{(r-1)(r-2)}(f_{\max,\Delta_n} - f_{\min,\Delta_n}). \tag{4.7}$$

**Proof.** We can write any cubic polynomial $f$ as

$$f = \sum_{i=1}^n f_i x_i^3 + \sum_{i<j}(f_{ij}x_i x_j^2 + g_{ij}x_i^2 x_j) + \sum_{i<j<k} f_{ijk}x_i x_j x_k.$$

Then, by (4.1) one can check that

$$\begin{aligned}
&\frac{(r-1)(r-2)}{r^2} f_{\min}^{(r-3)} \\
= \quad & \min_{\alpha\in I(n,r)} \left\{ f(\frac{\alpha}{r}) - \frac{1}{r^3}\left(3\sum_{i=1}^n f_i\alpha_i^2 - 2\sum_{i=1}^n f_i\alpha_i + \sum_{i<j}(f_{ij}+g_{ij})\alpha_i\alpha_j\right)\right\} \\
\geq \quad & f_{\Delta(n,r)} - \frac{1}{r}\max_{\alpha\in I(n,r)}\left\{3\sum_{i=1}^n f_i\left(\frac{\alpha_i}{r}\right)^2 + \sum_{i<j}(f_{ij}+g_{ij})\left(\frac{\alpha_i}{r}\right)\left(\frac{\alpha_j}{r}\right)\right\} \\
& +\frac{1}{r^2}\min_{\alpha\in I(n,r)} 2\sum_{i=1}^n f_i\frac{\alpha_i}{r} \\
\geq \quad & f_{\Delta(n,r)} - \frac{1}{r}\max_{x\in\Delta_n}\left\{3\sum_{i=1}^n f_i x_i^2 + \sum_{i<j}(f_{ij}+g_{ij})x_ix_j\right\} + \frac{1}{r^2}\min_{x\in\Delta_n} 2\sum_{i=1}^n f_ix_i. \quad (4.8)
\end{aligned}$$

Evaluating $f$ at $\mathbf{e}_i$ and $(\mathbf{e}_i+\mathbf{e}_j)/2$ yields, respectively, the relations:

$$f_{\min,\Delta_n} \leq f_i \leq f_{\max,\Delta_n}, \quad (4.9)$$

$$f_i + f_j + f_{ij} + g_{ij} \leq 8 f_{\max,\Delta_n}. \quad (4.10)$$

Using (4.10) and the fact that $\sum_{i=1}^n x_i = 1$, one can obtain

$$\sum_{i<j}(f_{ij}+g_{ij})x_ix_j \leq \sum_{i<j}(8f_{\max,\Delta_n} - f_i - f_j)x_ix_j = 8f_{\max,\Delta_n}\sum_{i<j}x_ix_j - \sum_{i=1}^n f_ix_i(1-x_i). \quad (4.11)$$

By (4.8), (4.9), (4.11) and the fact that $\sum_{i=1}^n x_i = 1$, one can get

$$\begin{aligned}
(r-1)(r-2)f_{\min}^{(r-3)} \quad & \geq \quad r^2 f_{\Delta(n,r)} - 4r f_{\max,\Delta_n} + (r+2)\min_{x\in\Delta_n}\sum_{i=1}^n f_ix_i \\
& \geq \quad r^2 f_{\Delta(n,r)} - 4r f_{\max,\Delta_n} + (r+2) f_{\min,\Delta_n}.
\end{aligned}$$

Hence, one has

$$\begin{aligned}
(r-1)(r-2)\left(f_{\Delta(n,r)} - f_{\min}^{(r-3)}\right) \quad & \leq \quad 4r f_{\max,\Delta_n} - (3r-2)f_{\Delta(n,r)} - (r+2) f_{\min,\Delta_n} \\
& \leq \quad 4r(f_{\max,\Delta_n} - f_{\min,\Delta_n}).
\end{aligned}$$

□

Now we observe that our result (4.7) refines the relevant upper bound obtained from [21, 22]. De Klerk et al. [21] show the following result.

**Theorem 4.6.** *[21, Theorem 3.3] Suppose $f \in \mathcal{H}_{n,3}$ and $r \geq 3$. Then*

$$f_{\min,\Delta_n} - f_{\min}^{(r-3)} \leq \frac{4r}{(r-1)(r-2)}(f_{\max,\Delta_n} - f_{\min,\Delta_n}), \tag{4.12}$$

$$f_{\Delta(n,r)} - f_{\min,\Delta_n} \leq \frac{4}{r}(f_{\max,\Delta_n} - f_{\min,\Delta_n}). \tag{4.13}$$

Recall in Corollary 2.12, we refine (4.13) to

$$f_{\Delta(n,r)} - f_{\min,\Delta_n} \leq \left(\frac{4}{r} - \frac{4}{r^2}\right)(f_{\max,\Delta_n} - f_{\min,\Delta_n}). \tag{4.14}$$

Similar to the quadratic case (in Section 4.2.1), our new upper bound (4.7) implies the upper bound obtained by adding up (4.12) and (4.14). However, we do not know any example showing that the upper bound (4.7) is tight. Thus, it is still an open question to show the tightness of the upper bound (4.7).

### 4.2.3 The square-free case

**Theorem 4.7.** *For any square-free polynomial $f = \sum_{I:I\subseteq[n],|I|=d} f_I x^I$ and $r \geq d$, one has*

$$f_{\Delta(n,r)} - f_{\min}^{(r-d)} \leq \left(\frac{r^d}{r^{\underline{d}}} - 1\right)(f_{\max,\Delta_n} - f_{\min,\Delta_n}). \tag{4.15}$$

**Proof.** From (4.1), one can easily check that

$$f_{\min}^{(r-d)} = \min_{\alpha\in I(n,r)} \sum_{I:I\subseteq[n],|I|=d} f_I \frac{\alpha^I}{r^{\underline{d}}} = \frac{1}{r^{\underline{d}}} \min_{\alpha\in I(n,r)} f(\alpha).$$

As a result, one can obtain

$$\frac{f_{\min}^{(r-d)}}{f_{\Delta(n,r)}} = \frac{r^d}{r^{\underline{d}}}.$$

For $d = 1$, the result (4.15) is clear.

Now we assume $d \geq 2$. Considering $f_{\max,\Delta_n} \geq 0$ (as $f(\mathbf{e}_i) = 0$ for any $i \in [n]$), we obtain

$$\begin{aligned} f_{\Delta(n,r)} - f_{\min}^{(r-d)} &= \left(1 - \frac{r^d}{r^{\underline{d}}}\right) f_{\Delta(n,r)} \leq \left(1 - \frac{r^d}{r^{\underline{d}}}\right) f_{\min,\Delta_n} \\ &\leq \left(\frac{r^d}{r^{\underline{d}}} - 1\right) (f_{\max,\Delta_n} - f_{\min,\Delta_n}). \end{aligned} \tag{4.16}$$

□

The following example shows that our upper bound (4.15) can be tight.

**Example 4.8.** *[Example 2.16 continued] Consider the square-free polynomial* $f = -x_1x_2$. *Recall that* $f_{\max,\Delta_n} = 0$, $f_{\min,\Delta_n} = -\frac{1}{4}$, *and*

$$f_{\Delta(2,r)} = \begin{cases} -\frac{1}{4} & \text{if } r \text{ is even,} \\ -\frac{1}{4} + \frac{1}{4r^2} & \text{if } r \text{ is odd.} \end{cases}$$

*By (4.1), we have*

$$f_{\Delta(2,r)} - f_{\min}^{(r-2)} = \begin{cases} \frac{1}{r-1} (f_{\max,\Delta_n} - f_{\min,\Delta_n}) & \text{if } r \text{ is even,} \\ \left(\frac{1}{r} + \frac{1}{r^2}\right) (f_{\max,\Delta_n} - f_{\min,\Delta_n}) & \text{if } r \text{ is odd.} \end{cases}$$

*For this example, the upper bound (4.15) is tight when* $r$ *is even. In fact, from (4.16), one can easily see that the upper bound (4.15) is tight as long as* $f_{\Delta(n,r)} = f_{\min,\Delta_n} - f_{\max,\Delta_n}$ *holds.*

### 4.2.4 The general case

We now consider an arbitrary polynomial $f \in \mathcal{H}_{n,d}$ and show the following upper bound for the range $f_{\Delta(n,r)} - f_{\min}^{(r-d)}$.

**Theorem 4.9.** *For any polynomial* $f \in \mathcal{H}_{n,d}$ *and* $r \geq d$, *one has*

$$f_{\Delta(n,r)} - f_{\min}^{(r-d)} \leq \frac{(r+d-1)^{\underline{d}} - r^d}{r^{\underline{d}}} \binom{2d-1}{d} d^d (f_{\max,\Delta_n} - f_{\min,\Delta_n}). \tag{4.17}$$

Note that when $f$ is quadratic, cubic or square-free, we have shown better upper bounds in Theorems 4.2, 4.5 and 4.7.

In the proof we will need the Vandermonde-Chu identity:

$$(\sum_{i=1}^{n} x_i)^{\underline{d}} = \sum_{\alpha \in I(n,d)} \frac{d!}{\alpha!} x^{\underline{\alpha}} \quad \forall x \in \mathbb{R}^n,$$

(recall (3.7)), which is an analogue of the multinomial theorem

$$(\sum_{i=1}^{n} x_i)^d = \sum_{\alpha\in I(n,d)} \frac{d!}{\alpha!} x^\alpha.$$

We will also need the following notation. For any univariate polynomial $t^d - t^{\underline{d}}$ (in the variable $t$), we can write it as

$$t^d - t^{\underline{d}} = \sum_{k=1}^{d-1} (-1)^{d-k-1} a_{d-k} t^k, \tag{4.18}$$

for some positive scalars $a_1, a_2, \ldots, a_{d-1}$. Then, one can easily check that

$$\sum_{k=1}^{d-1} a_{d-k} t^k = (t+d-1)^{\underline{d}} - t^d. \tag{4.19}$$

Analogously, for any $\beta \in I(n,d)$ and $x \in \mathbb{R}^n$, we can write the polynomial $x^\beta - x^{\underline{\beta}}$ as

$$x^\beta - x^{\underline{\beta}} = \sum_{\gamma:|\gamma|\le d-1} (-1)^{d-|\gamma|-1} c_\gamma^\beta x^\gamma, \tag{4.20}$$

for some nonnegative scalars $c_\gamma^\beta$.

Now we are ready to prove Theorem 4.9.

**Proof.** *(of Theorem 4.9)* From (4.1), we have

$$\frac{r^{\underline{d}}}{r^d} f_{\min}^{(r-d)} = \min_{\alpha\in I(n,r)} \left\{ \sum_{\beta\in I(n,d)} f_\beta \frac{\alpha^\beta}{r^d} - \sum_{\beta\in I(n,d)} f_\beta \frac{\alpha^\beta - \alpha^{\underline{\beta}}}{r^d} \right\}.$$

From this we obtain the inequality:

$$\frac{r^{\underline{d}}}{r^d} f_{\min}^{(r-d)} \ge f_{\Delta(n,r)} - \max_{\alpha\in I(n,r)} \sum_{\beta\in I(n,d)} f_\beta \frac{\alpha^\beta - \alpha^{\underline{\beta}}}{r^d}. \tag{4.21}$$

We now focus on the summation $\sum_{\beta\in I(n,d)} f_\beta(\alpha^\beta - \alpha^{\underline{\beta}})$.

We partition $[d-1]$ as $[d-1] = I_o \cup I_e$, where $I_o := \{k : k \in [d-1], d-k \text{ is odd}\}$ and $I_e := \{k : k \in [d-1], d-k \text{ is even}\}$. Then, from (4.20), the summation $\sum_{\beta \in I(n,d)} f_\beta(\alpha^\beta - \alpha^{\underline{\beta}})$ becomes

$$\begin{aligned}
\sum_{\beta \in I(n,d)} f_\beta(\alpha^\beta - \alpha^{\underline{\beta}}) &= \sum_{\beta \in I(n,d)} f_\beta \sum_{\gamma : |\gamma| \le d-1} (-1)^{d-|\gamma|-1} c_\gamma^\beta \alpha^\gamma \\
&= \sum_{k=1}^{d-1} \sum_{\gamma \in I(n,k)} \sum_{\beta \in I(n,d)} f_\beta (-1)^{d-|\gamma|-1} c_\gamma^\beta \alpha^\gamma \\
&\le \left( \max_{\beta \in I(n,d)} f_\beta \frac{\beta!}{d!} \right) \sum_{k \in I_o} \sum_{\gamma \in I(n,k)} \sum_{\beta \in I(n,d)} \frac{d!}{\beta!} c_\gamma^\beta \alpha^\gamma \\
&\quad - \left( \min_{\beta \in I(n,d)} f_\beta \frac{\beta!}{d!} \right) \sum_{k \in I_e} \sum_{\gamma \in I(n,k)} \sum_{\beta \in I(n,d)} \frac{d!}{\beta!} c_\gamma^\beta \alpha^\gamma. \quad (4.22)
\end{aligned}$$

Then, we consider the summation $\sum_{\gamma \in I(n,k)} \sum_{\beta \in I(n,d)} \frac{d!}{\beta!} c_\gamma^\beta \alpha^\gamma$ (for $k \in I_o$ or $I_e$) in the above inequality. We claim that, for any fixed $k \in [d-1]$, the following identity holds:

$$\sum_{\gamma \in I(n,k)} \sum_{\beta \in I(n,d)} \frac{d!}{\beta!} (-1)^{d-|\gamma|-1} c_\gamma^\beta x^\gamma = (-1)^{d-k-1} a_{d-k} (\sum_{i=1}^n x_i)^k, \quad (4.23)$$

where the constants $a_{d-k}$ are defined as in (4.18).

For this, observe that the polynomials at both sides of (4.23) are homogeneous of degree $k$. Hence (4.23) will follow if we can show that the equality holds after summing each side over $k \in [d-1]$. In other words, it suffices to show the identity:

$$\sum_{k=1}^{d-1} \sum_{\gamma \in I(n,k)} \sum_{\beta \in I(n,d)} \frac{d!}{\beta!} (-1)^{d-|\gamma|-1} c_\gamma^\beta x^\gamma = \sum_{k=1}^{d-1} (-1)^{d-k-1} a_{d-k} (\sum_{i=1}^n x_i)^k.$$

By the definition of $a_{d-k}$ in (4.18), the right side of the above equation is equal to $(\sum_{i=1}^n x_i)^d - (\sum_{i=1}^n x_i)^{\underline{d}}$. Hence, we only need to show

$$\sum_{k=1}^{d-1} \sum_{\gamma \in I(n,k)} \sum_{\beta \in I(n,d)} \frac{d!}{\beta!} (-1)^{d-|\gamma|-1} c_\gamma^\beta x^\gamma = (\sum_{i=1}^n x_i)^d - (\sum_{i=1}^n x_i)^{\underline{d}}. \quad (4.24)$$

Summing over (4.20), we obtain

$$\begin{aligned}\sum_{\beta\in I(n,d)}\frac{d!}{\beta!}\left(x^\beta - x^{\underline{\beta}}\right) &= \sum_{\beta\in I(n,d)}\sum_{\gamma:|\gamma|\le d-1}\frac{d!}{\beta!}(-1)^{d-|\gamma|-1}c_\gamma^\beta x^\gamma \\ &= \sum_{k=1}^{d-1}\sum_{\gamma\in I(n,k)}\sum_{\beta\in I(n,d)}\frac{d!}{\beta!}(-1)^{d-|\gamma|-1}c_\gamma^\beta x^\gamma.\end{aligned}$$

We can now conclude the proof of (4.24) (and thus of (4.23)). Indeed, by using the multinomial theorem and the Vandermonde-Chu identity (3.7), we see that the left-most side in the above relation is equal to $(\sum_{i=1}^n x_i)^d - (\sum_{i=1}^n x_i)^{\underline{d}}$.

Now we continue to analyze (4.22). By (4.23) we obtain

$$\begin{aligned}&\sum_{\beta\in I(n,d)} f_\beta(\alpha^\beta - \alpha^{\underline{\beta}}) \le \left(\max_{\beta\in I(n,d)} f_\beta\frac{\beta!}{d!}\right)\sum_{k\in I_o} a_{d-k}(\sum_{i=1}^n \alpha_i)^k \\ &-\left(\min_{\beta\in I(n,d)} f_\beta\frac{\beta!}{d!}\right)\sum_{k\in I_e} a_{d-k}(\sum_{i=1}^n \alpha_i)^k.\end{aligned}$$

Combining with (4.21), we get

$$r^{\underline{d}} f_{\min}^{(r-d)} \ \ge\ r^d f_{\Delta(n,r)} - \left(\max_{\beta\in I(n,d)} f_\beta\frac{\beta!}{d!}\right)\sum_{k\in I_o} a_{d-k}r^k + \left(\min_{\beta\in I(n,d)} f_\beta\frac{\beta!}{d!}\right)\sum_{k\in I_e} a_{d-k}r^k.$$

That is,

$$\begin{aligned}r^{\underline{d}}(f_{\Delta(n,r)} - f_{\min}^{(r-d)}) \ &\le\ (r^{\underline{d}} - r^d)f_{\Delta(n,r)} + \left(\max_{\beta\in I(n,d)} f_\beta\frac{\beta!}{d!}\right)\sum_{k\in I_o} a_{d-k}r^k \\ &\quad - \left(\min_{\beta\in I(n,d)} f_\beta\frac{\beta!}{d!}\right)\sum_{k\in I_e} a_{d-k}r^k.\end{aligned}$$

Since $r^{\underline{d}} - r^d = \sum_{k=1}^{d-1}(-1)^{d-k}a_{d-k}r^k$, we obtain

$$\begin{aligned} r^{\underline{d}}(f_{\Delta(n,r)} - f_{\min}^{(r-d)}) &\leq \sum_{k=1}^{d-1}(-1)^{d-k}a_{d-k}r^k f_{\Delta(n,r)} + \left(\max_{\beta\in I(n,d)} f_\beta \frac{\beta!}{d!}\right)\sum_{k\in I_o} a_{d-k}r^k \\ &\quad - \left(\min_{\beta\in I(n,d)} f_\beta \frac{\beta!}{d!}\right)\sum_{k\in I_e} a_{d-k}r^k \\ &= \left(\max_{\beta\in I(n,d)} f_\beta \frac{\beta!}{d!}\right)\sum_{k\in I_o} a_{d-k}r^k + f_{\Delta(n,r)}\sum_{k\in I_e} a_{d-k}r^k \\ &\quad - \left(\min_{\beta\in I(n,d)} f_\beta \frac{\beta!}{d!}\right)\sum_{k\in I_e} a_{d-k}r^k - f_{\Delta(n,r)}\sum_{k\in I_o} a_{d-k}r^k. \end{aligned}$$

According to (2.10), one has $\min_{\beta\in I(n,d)} f_\beta \frac{\beta!}{d!} \leq f_{\Delta(n,r)} \leq \max_{\beta\in I(n,d)} f_\beta \frac{\beta!}{d!}$. Therefore, we have

$$r^{\underline{d}}(f_{\Delta(n,r)} - f_{\min}^{(r-d)}) \leq \left(\max_{\beta\in I(n,d)} f_\beta \frac{\beta!}{d!} - \min_{\beta\in I(n,d)} f_\beta \frac{\beta!}{d!}\right)\sum_{k=1}^{d-1} a_{d-k}r^k.$$

That is,

$$f_{\Delta(n,r)} - f_{\min}^{(r-d)} \leq \frac{\sum_{k=1}^{d-1} a_{d-k}r^k}{r^{\underline{d}}}\left(\max_{\beta\in I(n,d)} f_\beta \frac{\beta!}{d!} - \min_{\beta\in I(n,d)} f_\beta \frac{\beta!}{d!}\right).$$

Finally, together with Theorem 2.6 and (4.19), we can conclude the result of Theorem 4.9. □

Now we compare the following theorem by De Klerk et al. [21] with our new result (4.17).

**Theorem 4.10.** *[21, Theorem 1.3] Suppose $f \in \mathcal{H}_{n,d}$ and $r \geq d$. Then*

$$f_{\min,\Delta_n} - f_{\min}^{(r-d)} \leq \left(\frac{r^d}{r^{\underline{d}}} - 1\right)\binom{2d-1}{d}d^d(f_{\max,\Delta_n} - f_{\min,\Delta_n}), \tag{4.25}$$

$$f_{\Delta(n,r)} - f_{\min,\Delta_n} \leq \left(1 - \frac{r^{\underline{d}}}{r^d}\right)\binom{2d-1}{d}d^d(f_{\max,\Delta_n} - f_{\min,\Delta_n}). \tag{4.26}$$

By adding up (4.25) and (4.26), we obtain

$$f_{\Delta(n,r)} - f_{\min}^{(r-d)} \leq \left(\frac{r^d}{r^{\underline{d}}} - \frac{r^{\underline{d}}}{r^d}\right)\binom{2d-1}{d}d^d(f_{\max,\Delta_n} - f_{\min,\Delta_n}). \tag{4.27}$$

**Lemma 4.11.** *When $r$ is large enough, the upper bound (4.17) refines the upper bound (4.27).*

**Proof.** It suffices to show that $\frac{r^d}{r^{\underline{d}}}-\frac{r^{\underline{d}}}{r^d}$ is larger than $\frac{\sum_{k=1}^{d-1} a_{d-k} r^k}{r^{\underline{d}}}$ when $r$ is sufficiently large. Since $\frac{r^d}{r^{\underline{d}}}-\frac{r^{\underline{d}}}{r^d} = (r^d - \frac{(r^{\underline{d}})^2}{r^d})/r^{\underline{d}}$, we only need to compare $r^d - \frac{(r^{\underline{d}})^2}{r^d}$ and $\sum_{k=1}^{d-1} a_{d-k} r^k$. For the term $r^d - \frac{(r^{\underline{d}})^2}{r^d}$, one can check that the coefficient of $r^d$ is 0 and the coefficient of $r^{d-1}$ is $2a_1 > 0$. On the other hand, in the summation $\sum_{k=1}^{d-1} a_{d-k} r^k$, the coefficient of $r^{d-1}$ is $a_1 > 0$. Therefore, when $r$ is sufficiently large, $r^d - \frac{(r^{\underline{d}})^2}{r^d}$ is larger than $\sum_{k=1}^{d-1} a_{d-k} r^k$, by which we conclude the proof. □

We now consider the following example, in which the upper bound (4.17) refines the upper bound (4.27) for $r \geq 10$.

**Example 4.12.** *Consider a polynomial $f \in \mathcal{H}_{n,4}$ written as*

$$\begin{aligned} f &= \sum_{i=1}^{n} f_i x_i^4 + \sum_{i<j} \left(f_{ij} x_i^3 x_j + g_{ij} x_i^2 x_j^2 + h_{ij} x_i x_j^3\right) + \sum_{i<j<k} \left(f_{ijk} x_i^2 x_j x_k\right. \\ &+ \left. g_{ijk} x_i x_j^2 x_k + h_{ijk} x_i x_j x_k^2\right) + \sum_{i<j<k<l} f_{ijkl} x_i x_j x_k x_l. \end{aligned}$$

*In this case, (4.17) reads*

$$f_{\Delta(n,r)} - f_{\min}^{(r-4)} \leq \frac{6r^2+11r+6}{(r-1)(r-2)(r-3)} \binom{7}{4} 4^4 (f_{\max,\Delta_n} - f_{\min,\Delta_n}), \tag{4.28}$$

*while (4.27) reads*

$$f_{\Delta(n,r)} - f_{\min}^{(r-4)} \leq \frac{12r^2 - 58r + 144 - \frac{193}{r} + \frac{132}{r^2} - \frac{36}{r^3}}{(r-1)(r-2)(r-3)} \binom{7}{4} 4^4 (f_{\max,\Delta_n} - f_{\min,\Delta_n}). \tag{4.29}$$

*One can check that (4.28) refines (4.29) when $r \geq 10$.*

## 4.3 Concluding remarks

We now consider the convergence rate of the sequence

$$\alpha_r(f) := \frac{f_{\Delta(n,r)} - f_{\min}^{(r-d)}}{f_{\max,\Delta_n} - f_{\min,\Delta_n}} \qquad r = 1, 2, \ldots$$

Note that our results in Theorems 4.2, 4.5, 4.7 and 4.9 imply that

$$\alpha_r(f) \leq \frac{C(d)}{r} = O\left(\frac{1}{r}\right),$$

where $C(d)$ is a constant depending on $d$.

Moreover, from Example 4.8, we can conclude that the above dependence on $r$ of $\alpha_r(f)$ is tight, in the sense that there does not exist any $\epsilon > 0$ such that $\alpha_r(f) = O(1/r^{1+\epsilon})$.

# Part II

# Polynomial Optimization over a Compact Set

# Chapter 5

# Lasserre's measure-based hierarchy of upper bounds

## 5.1 Introduction

In this part we consider the problem of minimizing a continuous function $f : \mathbb{R}^n \to \mathbb{R}$ over a compact set $\mathbf{K} \subseteq \mathbb{R}^n$. That is, we consider the problem of computing the parameter:

$$f_{\min,\mathbf{K}} = \min_{x \in \mathbf{K}} f(x). \tag{5.1}$$

This problem contains polynomial optimization over the standard simplex (1.21) (studied in Part I) as a special case, and thus is NP-hard in general.

As we saw in (1.14), one can reformulate the optimization problem (5.1) as the problem of finding a probability measure $\mu$ on $\mathbf{K}$ minimizing the integral $\int_{\mathbf{K}} f(x)\mu(dx)$, i.e.,

$$f_{\min,\mathbf{K}} = \min_{\mu \in \mathcal{M}(\mathbf{K})} \int_{\mathbf{K}} f(x)\mu(dx). \tag{5.2}$$

By selecting suitable probability measures on $\mathbf{K}$, one obtains upper bounds for $f_{\min,\mathbf{K}}$. In Chapter 2 and Chapter 3 this approach has been investigated, in particular, to construct upper bounds for the parameter $f_{\Delta(n,r)}$ by selecting some discrete distributions over $\Delta(n, r)$.

In this chapter, we investigate a hierarchy of upper bounds for $f_{\min,\mathbf{K}}$ proposed by Lasserre [53]. More precisely, Lasserre [53] shows that, in order to compute $f_{\min,\mathbf{K}}$, it suffices to search for a sums of squares density function $h$ which minimizes the integral $\int_{\mathbf{K}} fh dx$, see Theorem 5.1 below. Then by adding degree constraints on the polynomial density $h$, one can obtain the following upper bound $\underline{f}_{\mathbf{K}}^{(r)}$ for $f_{\min,\mathbf{K}}$, which was defined in (1.16):

$$\underline{f}_{\mathbf{K}}^{(r)} = \inf_{h\in\Sigma[x]_r} \int_{\mathbf{K}} h(x)f(x)dx \ \text{ s.t. } \int_{\mathbf{K}} h(x)dx = 1.$$

In this chapter we analyze the quality of $\underline{f}_{\mathbf{K}}^{(r)}$. Our main result is Theorem 5.7, which states that, under some conditions on $f$ and $\mathbf{K}$, the convergence rate of $\underline{f}_{\mathbf{K}}^{(r)}$ satisfies

$$\underline{f}_{\mathbf{K}}^{(r)} - f_{\min,\mathbf{K}} \leq \frac{\zeta(\mathbf{K})M_f}{\sqrt{r}} = O\left(\frac{1}{\sqrt{r}}\right),$$

where $\zeta(\mathbf{K})$ is a constant depending only on $\mathbf{K}$ and $M_f$ is the Lipschitz constant of $f$ on $\mathbf{K}$. Namely, this result holds when $f$ is Lipschitz continuous and $\mathbf{K}$ is a full-dimentional compact set satisfying an additional geometrical condition, see Assumption 5.4 below.

In addition, as an application, we indicate how to sample feasible points in $\mathbf{K}$ based on the optimal density function $h^*$ in the problem of computing $\underline{f}_{\mathbf{K}}^{(r)}$. In several numerical examples, these points perform better than the points obtained by uniform sampling from $\mathbf{K}$. See Sections 5.4 and 5.5 for details.

### 5.1.1 Lasserre's hierarchy of upper bounds

Now we recall the result of Lasserre [53], which roughly speaking says that, in (5.2), we may restrict to the Lebesgue measure with an arbitrary sum of squares of polynomials density function.

**Theorem 5.1.** *[53, Theorem 3.2] Let $\mathbf{K} \subseteq \mathbb{R}^n$ be compact and let $f$ be a continuous function on $\mathbb{R}^n$. Then the minimum of $f$ over $\mathbf{K}$ can be expressed as*

$$f_{\min,\mathbf{K}} = \inf_{h\in\Sigma[x]} \int_{\mathbf{K}} h(x)f(x)dx \quad s.t. \int_{\mathbf{K}} h(x)dx = 1. \tag{5.3}$$

As recalled above, by bounding the degree of the polynomial $h \in \Sigma[x]$ by $2r$ in (5.3), we obtain the parameter:

$$\underline{f}_{\mathbf{K}}^{(r)} = \inf_{h\in\Sigma[x]_r} \int_{\mathbf{K}} h(x)f(x)dx \ \text{ s.t. } \int_{\mathbf{K}} h(x)dx = 1. \tag{5.4}$$

Clearly, the following inequalities hold for any integer $r \geq 1$:

$$f_{\min,\mathbf{K}} \leq \cdots \leq \underline{f}_{\mathbf{K}}^{(r+1)} \leq \underline{f}_{\mathbf{K}}^{(r)} \leq \cdots \leq \underline{f}_{\mathbf{K}}^{(1)}.$$

Additionally, Lasserre [53] gives conditions under which the infimum in (5.4) is attained.

**Theorem 5.2.** *[53, Theorems 4.1 and 4.2] Assume $\mathbf{K} \subseteq \mathbb{R}^n$ is compact and has nonempty interior and let $f$ be a polynomial. Then, the program (5.4) has an optimal solution for every $r \in \mathbb{N}$ and*

$$\lim_{r\to\infty} \underline{f}_{\mathbf{K}}^{(r)} = f_{\min,\mathbf{K}}.$$

When $f$ is a polynomial, we now recall how to compute the parameter $\underline{f}_{\mathbf{K}}^{(r)}$ in terms of the moments $m_\alpha(\mathbf{K})$ of the Lebesgue measure on $\mathbf{K}$, where

$$m_\alpha(\mathbf{K}) := \int_{\mathbf{K}} x^\alpha dx \quad \text{ for } \alpha \in \mathbb{N}^n.$$

Recall $N(n,r) = \{\alpha \in \mathbb{N}^n : \sum_{i=1}^n \alpha_i \leq r\}$, and suppose $f(x) = \sum_{\beta\in N(n,d)} f_\beta x^\beta$ has degree $d$. If we write $h \in \Sigma[x]_r$ as $h(x) = \sum_{\alpha\in N(n,2r)} h_\alpha x^\alpha$, then $\underline{f}_{\mathbf{K}}^{(r)}$ can be reformulated as follows:

$$\begin{aligned} \underline{f}_{\mathbf{K}}^{(r)} \;=\; \min & \sum_{\beta\in N(n,d)} f_\beta \sum_{\alpha\in N(n,2r)} h_\alpha m_{\alpha+\beta}(\mathbf{K}) \\ \text{s.t.} & \sum_{\alpha\in N(n,2r)} h_\alpha m_\alpha(\mathbf{K}) = 1, \\ & \sum_{\alpha\in N(n,2r)} h_\alpha x^\alpha \in \Sigma[x]_r. \end{aligned} \tag{5.5}$$

Hence, if we know the moments $m_\alpha(\mathbf{K})$ for any $\alpha \in \mathbb{N}^n$ with $|\alpha| = \sum_{i=1}^n \alpha_i \leq d + 2r$, then we can compute the parameter $\underline{f}_{\mathbf{K}}^{(r)}$ by solving a semidefinite program, which involves a positive semidefinite matrix of size $\binom{n+r}{r}$.
In particular, when $\mathbf{K}$ is the (full-dimensional) simplex

$$\widehat{\Delta_n} = \{x \in \mathbb{R}^n_+ : \sum_{i=1}^n x_i \leq 1\},$$

the unit hypercube
$$\mathbf{Q}_n = [0,1]^n,$$
or the unit ball
$$\mathbf{B}_1(0) = \{x \in \mathbb{R}^n : \|x\| \leq 1\},$$
there exist explicit formulas for the moments $m_\alpha(\mathbf{K})$, as recalled below.

**Proposition 5.3.** *When $\mathbf{K}$ is $\widehat{\Delta_n}$, $\mathbf{Q}_n$, or $\mathbf{B}_1(0)$, the moment $m_\alpha(\mathbf{K})$ for any $\alpha \in \mathbb{N}^n$ reads as follows:*
$$m_\alpha(\widehat{\Delta_n}) = \frac{\prod_{i=1}^n \alpha_i!}{(|\alpha|+n)!}, \tag{5.6}$$

$$m_\alpha(\mathbf{Q}_n) = \prod_{i=1}^n \frac{1}{\alpha_i+1}, \tag{5.7}$$

*and*

$$m_\alpha(\mathbf{B}_1(0)) = \begin{cases} \frac{\pi^{(n-1)/2} 2^{(n+1)/2} \prod_{i=1}^n (\alpha_i - 1)!!}{(n+|\alpha|)!!} & \text{if } \alpha_i \text{ is even for all } i \in [n], \\ 0 & \text{otherwise.} \end{cases} \tag{5.8}$$

**Proof.** For (5.6), see, e.g., [48, equation (2.4)] or [37, equation (2.2)].
By direct calculation, one can obtain (5.7):

$$m_\alpha(\mathbf{Q}_n) = \int_{\mathbf{Q}_n} x^\alpha dx = \prod_{i=1}^n \int_0^1 x_i^{\alpha_i} dx_i = \prod_{i=1}^n \frac{1}{\alpha_i+1}.$$

One may prove relation (5.8) by combining

$$\int_{\mathbf{B}_1(0)} x^\alpha dx = \frac{1}{\Gamma(1+(n+|\alpha|)/2)} \int_{\mathbb{R}^n} x^\alpha \exp\left(-\|x\|^2\right) dx$$

(see, e.g. [54, Theorem 2.1]), and the fact that

$$\int_{-\infty}^{+\infty} t^p \exp\left(-t^2/2\right) dt = \begin{cases} \sqrt{2\pi}(p-1)!! & \text{if } p \text{ is even,} \\ 0 & \text{if } p \text{ is odd,} \end{cases}$$

(see, e.g., page 872 in [53]), together with the identity $\Gamma(1+\frac{k}{2}) = \frac{k!!}{2^{(k+1)/2}}\sqrt{\pi}$ for any integer $k \in \mathbb{N}$ (see, e.g., [1, Section 6.1.12]). □

However, for a general polytope $\mathbf{K} \subseteq \mathbb{R}^n$, it is a hard problem to compute the moments $m_\alpha(\mathbf{K})$. In fact, the problem of computing the volume of polytopes of varying dimensions is already #P-hard [32]. On the other hand, any polytope $\mathbf{K}$ in $\mathbb{R}^n$ can be triangulated into finitely many simplices (see, e.g., [27]) so that one could use (5.6) to obtain the moments $m_\alpha(\mathbf{K})$ of $\mathbf{K}$. The complexity of this method depends on the number of simplices in the triangulation. However, this number can be exponentially large (e.g., for the hypercube) and the problem of finding the smallest possible triangulation of a polytope is NP-hard, even in fixed dimension $n = 3$ (see, e.g., [27]).

#### Example

Consider the minimization of the Motzkin polynomial

$$f(x_1, x_2) = x_1^4 x_2^2 + x_1^2 x_2^4 - 3x_1^2 x_2^2 + 1$$

over the hypercube $\mathbf{K} = [-2, 2]^2$, which has four global minimizers at the points $(\pm 1, \pm 1)$, and $f_{\min,\mathbf{K}} = 0$. Figure 5.1 shows the computed optimal sum of squares density function $h^*$, for $r = 8$, 10, and 12, corresponding to $\underline{f}_{\mathbf{K}}^{(8)} = 0.565553$, $\underline{f}_{\mathbf{K}}^{(10)} = 0.507829$, and $\underline{f}_{\mathbf{K}}^{(12)} = 0.406076$, respectively. We observe that the optimal density $h^*$ shows four peaks at the four global minimizers. Thus it appears that $h^*$ approximates the density of a convex combination of the Dirac measures at the four minimizers, where the 'variance' decreases as $r$ increases.
We will present several other numerical examples in Section 5.5.

### 5.1.2 Our main result

Our main result in this chapter is an upper bound for the range $\underline{f}_{\mathbf{K}}^{(r)} - f_{\min,\mathbf{K}}$, which applies to the case when $f$ is Lipschitz continuous on $\mathbf{K}$ and when $\mathbf{K}$ is a full-dimensional compact set satisfying the additional condition from Assumption 5.4, see Theorem 5.7 below.

In what follows, we first review background material on Lipschitz continuous functions, then we give the conditions for $\mathbf{K}$ needed for our result, and finally we state our main result in Theorem 5.7.

#### Lipschitz continuous functions

A function $f$ is said to be Lipschitz continuous on $\mathbf{K}$, with Lipschitz constant $M_f$, if it satisfies:

$$|f(y) - f(x)| \leq M_f \|y - x\| \quad \text{for all } x, y \in \mathbf{K}.$$

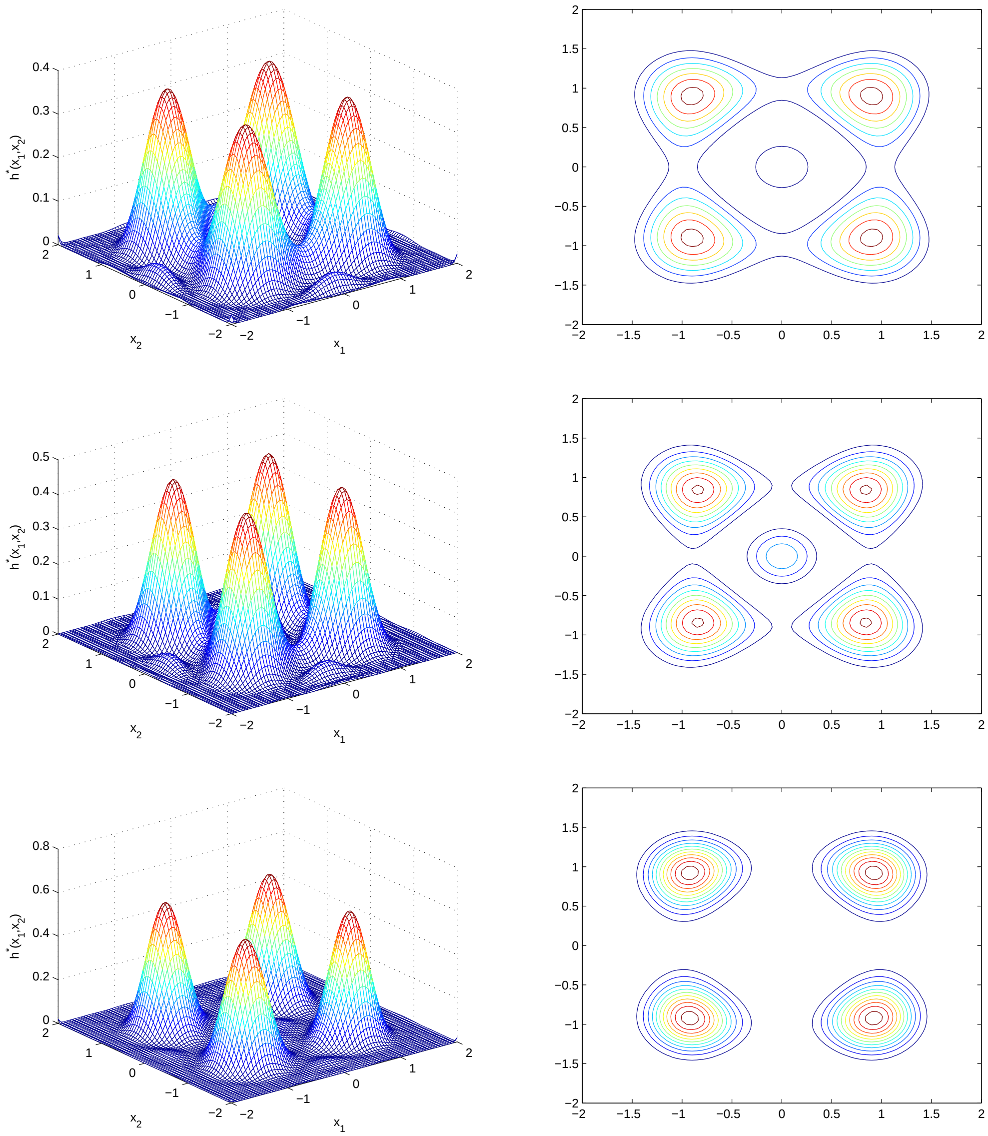


Figure 5.1: Graphs and contour plots of $h^*(x)$ on $[-2,2]^2$ ($r = 8, 10, 12$) for the Motzkin polynomial.

If $f$ is continuous and differentiable on $\mathbf{K}$, then $f$ is Lipschitz continuous on $\mathbf{K}$ with respect to the constant

$$M_f = \max_{x \in \mathbf{K}} \|\nabla f(x)\|. \tag{5.9}$$

Recall that $w_{\min}(\mathbf{K})$ denotes the minimal width of $\mathbf{K}$, which is defined as the minimum distance between two distinct parallel supporting hyperplanes of $\mathbf{K}$. Then, given any $n$-variate polynomial $f$ of degree $d$, the Markov inequality for $f$ on a convex body $\mathbf{K}$ reads as

$$\max_{x \in \mathbf{K}} \|\nabla f(x)\| \leq \frac{2d^2}{w_{\min}(\mathbf{K})} \sup_{x \in \mathbf{K}} |f(x)|,$$

see, e.g., [17, relation (8)]. Thus, together with (5.9), we have that $f$ is Lipschitz continuous on $\mathbf{K}$ with respect to the constant

$$M_f \leq \frac{2d^2}{w_{\min}(\mathbf{K})} \sup_{x \in \mathbf{K}} |f(x)|. \tag{5.10}$$

**Assumption on the set K**

Before giving the conditions for $\mathbf{K}$ in Assumption 5.4, we recall some notation.

Recall that $D(\mathbf{K}) = \max_{x,y \in \mathbf{K}} \|x - y\|^2$ denotes the (squared) diameter of the set $\mathbf{K}$, where $\|x\| = \sqrt{\sum_{i=1}^n x_i{}^2}$ is the $\ell_2$-norm. Moreover, $\mathbf{B}_\epsilon(a) = \{x \in \mathbb{R}^n : \|x - a\| \leq \epsilon\}$ denotes the Euclidean ball centered at $a \in \mathbb{R}^n$ and with radius $\epsilon > 0$. With $\gamma_n$ denoting the volume of the $n$-dimensional unit ball, the volume of the ball $\mathbf{B}_\epsilon(a)$ is given by $\mathrm{vol}\mathbf{B}_\epsilon(a) = \epsilon^n \gamma_n$.

Using the above notation, we can state our main assumption about the set $\mathbf{K}$ as follows.

**Assumption 5.4.** *There exist constants $\eta_{\mathbf{K}} > 0$ and $\epsilon_{\mathbf{K}} > 0$ such that, for any point $a \in \mathbf{K}$,*

$$\mathrm{vol}(\mathbf{B}_\epsilon(a) \cap \mathbf{K}) \geq \eta_{\mathbf{K}} \mathrm{vol}\mathbf{B}_\epsilon(a) = \eta_{\mathbf{K}} \epsilon^n \gamma_n, \quad \textit{for all } 0 < \epsilon \leq \epsilon_{\mathbf{K}}. \tag{5.11}$$

Roughly speaking, Assumption 5.4 requires that, at any point $a \in \mathbf{K}$, there is a ball centered at $a$, whose intersection with $\mathbf{K}$ is at least a constant fraction of the full ball.

In Section 5.3 below we will revisit Assumption 5.4. More precisely, we will consider its link to a condition classically used in approximation theory, known as the *interior cone condition* (see Definition 5.16 below). We will show that any set satisfying the interior cone condition also satisfies Assumption 5.4 (see Lemma 5.20 below). We will also show some instances that satisfy the interior cone condition (and thus Assumption 5.4), namely, full-dimensional bounded convex sets.

Moreover, for any compact set $\mathbf{K} \subseteq \mathbb{R}^n$ satisfying Assumption 5.4, define

$$r_{\mathbf{K}} := \max\left\{\frac{D(\mathbf{K})e}{2\epsilon_{\mathbf{K}}^3}, n\right\} \quad \text{if } \epsilon_{\mathbf{K}} \leq 1 \quad \text{and} \quad r_{\mathbf{K}} := \frac{D(\mathbf{K})e}{2} \quad \text{if } \epsilon_{\mathbf{K}} \geq 1. \tag{5.12}$$

**Remark 5.5.** *In fact, Assumption 5.4 involves some concepts that are closely related to the Lebesgue density. Indeed, given $\epsilon > 0$, the approximate density of $\mathbf{K}$ in an $\epsilon$-neighborhood of a point $a$ in $\mathbf{K}$ is defined as*

$$d_\epsilon(a) = \frac{\mathrm{vol}(\mathbf{K} \cap \mathbf{B}_\epsilon(a))}{\mathrm{vol}(\mathbf{B}_\epsilon(a))},$$

*and the Lebesgue density of $\mathbf{K}$ in $a$ is then defined as the limit of $d_\epsilon(a)$ as $\epsilon$ tends to $0$. (See, e.g., [68] for more information on the Lebesgue density).*

*Then, Assumption 5.4 can be stated as follows: there exist constants $\eta_{\mathbf{K}} > 0$ and $\epsilon_{\mathbf{K}} > 0$ such that, for any point $a$ in $\mathbf{K}$, the approximate density $d_\epsilon(a)$ is at least $\eta_{\mathbf{K}}$ for any $0 < \epsilon \leq \epsilon_{\mathbf{K}}$. Hence, Assumption 5.4 implies that the Lebesgue density of $\mathbf{K}$ in any point $a$ in $\mathbf{K}$ is at least $\eta_{\mathbf{K}}$.*

*We observe that our main result in Theorem 5.7 would in fact still hold under the weaker assumption that there exists a global minimizer $a$ in $\mathbf{K}$ at which the Lebesgue density of $\mathbf{K}$ is strictly positive; see Remark 5.15 for more details.*

**Example 5.6.** *We now give an example of a set $\mathbf{K}$ that does not satisfy Assumption 5.4. Consider the set $\mathbf{K} \subseteq \mathbb{R}^2$ in Figure 5.2, which can be described as:*

$$\mathbf{K} = \{x \in \mathbb{R}^2 \ : \ x \geq 0, (x_1 - 1)^2 + (x_2 - 1)^2 \geq 1\}.$$

*Observe that $\mathbf{K}$ is bounded by an arc and the two tangent lines. Then one can easily check that the constraint (5.11) does not hold for the two points $a$ and $b$.*

### The main result

We can now present our main result.

**Theorem 5.7.** *Assume that $\mathbf{K} \subseteq \mathbb{R}^n$ is compact and satisfies Assumption 5.4, and consider the parameter $r_{\mathbf{K}}$ from (5.12). Then there exists a constant $\zeta(\mathbf{K})$ (depending only on $\mathbf{K}$) such that, for any Lipschitz continuous function $f$ with Lipschitz constant $M_f$ on $\mathbf{K}$, the following inequality holds:*

$$\underline{f}_{\mathbf{K}}^{(r)} - f_{\min,\mathbf{K}} \leq \frac{\zeta(\mathbf{K}) M_f}{\sqrt{r}} \quad \textit{for any } r \geq r_{\mathbf{K}} + 1. \tag{5.13}$$

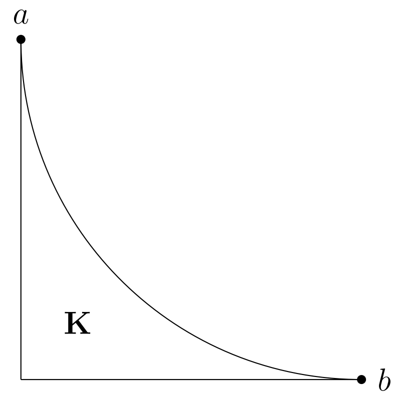


Figure 5.2: A set $\mathbf{K}$ that does not satisfy Assumption 5.4

*Moreover, if $f$ is a polynomial of degree $d$ and $\mathbf{K}$ is a convex body, then*

$$\underline{f}_{\mathbf{K}}^{(r)} - f_{\min,\mathbf{K}} \leq \frac{2d^2\zeta(\mathbf{K})\sup_{x\in\mathbf{K}}|f(x)|}{w_{\min}(\mathbf{K})\sqrt{r}} \quad \textit{for any } r \geq r_{\mathbf{K}} + 1. \tag{5.14}$$

The key idea to show this result is to select suitable sums of squares densities which we are able to analyse. For this, we will select a global minimizer $a$ of $f$ over $\mathbf{K}$ and consider the Gaussian distribution with mean $a$ and, as sums of squares densities, we will select the polynomials $H_{r,a}$ obtained by truncating the Taylor series expansion of the Gaussian distribution, see relation (5.17).

## 5.2 Proof for the convergence rate

In this section we prove our main result in Theorem 5.7.

In the first step we indicate in Section 5.2.1 how to select the polynomial density function $h$ as a special sum of squares that we will be able to analyze. Namely, we let $a$ denote a global minimizer of the function $f$ over the set $\mathbf{K} \subseteq \mathbb{R}^n$. Then we consider the density function $G_a$ in (5.15) of the Gaussian distribution with mean $a$ and the polynomial $H_{r,a}$ in (5.17), which is obtained from the truncation at degree $2r$ of the Taylor series expansion of the Gaussian density function $G_a$.

The second step will be to analyze the quality of the bound obtained by selecting the polynomial $H_{r,a}$ and this will be the most technical part of the proof, carried out in Section 5.2.2.

### 5.2.1 Choosing the polynomial density function $H_{r,a}$

Throughout this section, we let $a$ be a global minimizer of $f$ over $\mathbf{K}$. Consider the function

$$G_a(x) := \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{\|x-a\|^2}{2\sigma^2}\right), \tag{5.15}$$

which is the probability density function of the Gaussian distribution with mean $a$ and standard variance $\sigma$ (whose value will be defined later). Let the constant $C_{\mathbf{K},a}$ be defined by

$$\int_{\mathbf{K}} C_{\mathbf{K},a} G_a(x) dx = 1. \tag{5.16}$$

Observe that $G_a(x)$ is equal to the function $\frac{1}{(2\pi\sigma^2)^{n/2}} e^{-t}$ evaluated at the point $t = \frac{\|x-a\|^2}{2\sigma^2}$.

Denote by $H_{r,a}$ the Taylor series expansion of $G_a$ truncated at the order $2r$. That is,

$$H_{r,a}(x) = \frac{1}{(2\pi\sigma^2)^{n/2}} \sum_{k=0}^{2r} \frac{1}{k!} \left(-\frac{\|x-a\|^2}{2\sigma^2}\right)^k. \tag{5.17}$$

Moreover consider the constant $c^r_{\mathbf{K},a}$, defined by

$$\int_{\mathbf{K}} c^r_{\mathbf{K},a} H_{r,a}(x) dx = 1. \tag{5.18}$$

The next step is to show that $H_{r,a}$ is a sum of squares of polynomials and thus $H_{r,a} \in \Sigma[x]_{2r}$. This follows from the next lemma.

**Lemma 5.8.** *Let $\phi_{2r}(t)$ denote the (univariate) polynomial of degree $2r$ obtained by truncating the Taylor series expansion of $e^{-t}$ at the order $2r$. That is,*

$$\phi_{2r}(t) := \sum_{k=0}^{2r} \frac{(-t)^k}{k!}.$$

*Then $\phi_{2r}$ is a sum of squares of polynomials. Moreover, we have*

$$0 \leq \phi_{2r}(t) - e^{-t} \leq \frac{t^{2r+1}}{(2r+1)!} \quad \text{for all } t \geq 0. \tag{5.19}$$

**Proof.** First, we show that $\phi_{2r}$ is a sum of squares. As $\phi_{2r}$ is a univariate polynomial, by Hilbert's Theorem (see, e.g., [58, Theorem 3.4]), it suffices to show that $\phi_{2r}(t) \geq 0$ for any $t \in \mathbb{R}$. As $\phi_{2r}(-\infty) = \phi_{2r}(+\infty) = +\infty$, it suffices to show that $\phi_{2r}(t) \geq 0$ at all the stationary points $t$ where $\phi'_{2r}(t) = 0$. For this, observe that $\phi'_{2r}(t) = \sum_{k=1}^{2r}(-1)^k \frac{t^{k-1}}{(k-1)!}$, so that it can be written as $\phi'_{2r}(t) = -\phi_{2r}(t) + \frac{t^{2r}}{(2r)!}$. Hence, for any $t$ with $\phi'_{2r}(t) = 0$, we have $\phi_{2r}(t) = \frac{t^{2r}}{(2r)!} \geq 0$.

Next, we show that $\phi_{2r}(t) \geq e^{-t}$ for all $t \geq 0$. Fix $t \geq 0$. Then, by Taylor Theorem (see, e.g., [99]), one has $e^{-t} = \phi_{2r}(t) + \frac{\phi^{(2r+1)}(\xi)t^{2r+1}}{(2r+1)!}$ for some $\xi \in [0,t]$. As $\phi^{(2r+1)}(\xi) = -e^{-\xi}$, one can conclude that $e^{-t} - \phi_{2r}(t) = -\frac{e^{-\xi}t^{2r+1}}{(2r+1)!} \leq 0$ and $e^{-t} - \phi_{2r}(t) \geq -\frac{t^{2r+1}}{(2r+1)!}$. □

We now consider the parameter $f^{(r)}_{\mathbf{K},a}$ defined as

$$f^{(r)}_{\mathbf{K},a} := \int_{\mathbf{K}} f(x) c^r_{\mathbf{K},a} H_{r,a}(x) dx, \tag{5.20}$$

where $H_{r,a}(x)$ and $c^r_{\mathbf{K},a}$ are defined in (5.17) and (5.18).

Our main technical result is the following upper bound for the range $f^{(r)}_{\mathbf{K},a} - f_{\min,\mathbf{K}}$, whose proof is given in Section 5.2.2 below. Theorem 5.7 follows then as a direct application of Theorem 5.9.

**Theorem 5.9.** *Assume $\mathbf{K} \subseteq \mathbb{R}^n$ is compact and satisfies Assumption 5.4. Consider the parameter $f^{(r)}_{\mathbf{K},a}$ from* (5.20) *and the parameter $r_{\mathbf{K}}$ from (5.12). Then there exists a constant $\zeta(\mathbf{K})$ (depending only on $\mathbf{K}$) such that, for any Lipschitz continuous function $f$ with Lipschitz constant $M_f$ on $\mathbf{K}$, the following inequality holds:*

$$f^{(r)}_{\mathbf{K},a} - f_{\min,\mathbf{K}} \leq \frac{\zeta(\mathbf{K})M_f}{\sqrt{2r+1}}, \quad \textit{for any } r \geq \frac{r_{\mathbf{K}}}{2}. \tag{5.21}$$

*Moreover, if $f$ is a polynomial of degree $d$ and $\mathbf{K}$ is a convex body, then*

$$f^{(r)}_{\mathbf{K},a} - f_{\min,\mathbf{K}} \leq \frac{2d^2\zeta(\mathbf{K})\sup_{x\in\mathbf{K}}|f(x)|}{w_{\min}(\mathbf{K})\sqrt{2r+1}}, \quad \textit{for any } r \geq \frac{r_{\mathbf{K}}}{2}. \tag{5.22}$$

**Proof.** *(of Theorem 5.7)* Assume $f$ is Lipschitz continuous with Lipschitz constant $M_f$ on $K$ and $a$ is a minimizer of $f$ over the set $\mathbf{K}$. Using the definitions (5.4) and (5.20) of the parameters and the fact that $H_{r,a}$ is a sum of squares with degree $4r$, it follows that

$$\underline{f}^{(2r+1)}_{\mathbf{K}} \leq \underline{f}^{(2r)}_{\mathbf{K}} \leq f^{(r)}_{\mathbf{K},a}, \quad \text{for any } r \in \mathbb{N}.$$

Then, from inequality (5.21) in Theorem 5.9, one obtains

$$\underline{f}_{\mathbf{K}}^{(2r+1)} - f_{\min,\mathbf{K}} \leq \underline{f}_{\mathbf{K}}^{(2r)} - f_{\min,\mathbf{K}} \leq f_{\mathbf{K},a}^{(r)} - f_{\min,\mathbf{K}} \leq \frac{\zeta(\mathbf{K})M_f}{\sqrt{2r+1}} \quad \text{for any } r \geq \frac{r_{\mathbf{K}}}{2}.$$

Hence, for any $r \geq r_{\mathbf{K}} + 1$,

$$\begin{aligned} \underline{f}_{\mathbf{K}}^{(r)} - f_{\min,\mathbf{K}} &\leq \frac{\zeta(\mathbf{K})M_f}{\sqrt{r+1}} \leq \frac{\zeta(\mathbf{K})M_f}{\sqrt{r}} \quad \text{for even } r, \\ \underline{f}_{\mathbf{K}}^{(r)} - f_{\min,\mathbf{K}} &\leq \frac{\zeta(\mathbf{K})M_f}{\sqrt{r}} \quad \text{for odd } r. \end{aligned}$$

This concludes the proof for relation (5.13), and relation (5.14) follows from (5.22) in an analogous way. This finishes the proof of Theorem 5.7. □

### 5.2.2 Analyzing the polynomial density function $H_{r,a}$

In this subsection we prove the result of Theorem 5.9. Recall that $a$ is a global minimizer of $f$ over $\mathbf{K}$. For the proof, we will need the following four technical lemmas (Lemmas 5.10, 5.11, 5.12 and 5.13).

**Lemma 5.10.** *Assume $\mathbf{K} \subseteq \mathbb{R}^n$ is compact and satisfies Assumption 5.4. Then, for any $0 < \epsilon \leq \epsilon_{\mathbf{K}}$ and $r \in \mathbb{N}$, we have:*

$$c_{\mathbf{K},a}^r \leq C_{\mathbf{K},a} \leq \frac{(2\pi\sigma^2)^{n/2} \exp\left(\frac{\epsilon^2}{2\sigma^2}\right)}{\eta_{\mathbf{K}} \epsilon^n \gamma_n}. \tag{5.23}$$

**Proof.** By Lemma 5.8, $\phi_{2r}(t) \geq e^{-t}$ for all $t \geq 0$, which implies $H_{r,a}(x) \geq G_a(x)$ for all $x \in \mathbb{R}^n$. Together with the relations (5.16) and (5.18) defining the constants $C_{\mathbf{K},a}$ and $c_{\mathbf{K},a}^r$, we deduce that $c_{\mathbf{K},a}^r \leq C_{\mathbf{K},a}$. Moreover, by the definition (5.16) of the constant $C_{K,a}$, one has

$$\begin{aligned} \frac{1}{C_{\mathbf{K},a}} &= \int_{\mathbf{K}} G_a(x)dx = \int_{\mathbf{K}} \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{\|x-a\|^2}{2\sigma^2}\right) dx \\ &\geq \int_{\mathbf{K}\cap B_\epsilon(a)} \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{\|x-a\|^2}{2\sigma^2}\right) dx \\ &\geq \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{\epsilon^2}{2\sigma^2}\right) \mathrm{vol}(\mathbf{K} \cap \mathbf{B}_\epsilon(a)). \end{aligned}$$

We now use relation (5.11) from Assumption 5.4 in order to conclude that

$$\mathrm{vol}(\mathbf{K} \cap \mathbf{B}_\epsilon(a)) \geq \eta_{\mathbf{K}} \epsilon^n \gamma_n,$$

which gives the desired upper bound on $C_{K,a}$. ☐

**Lemma 5.11.** *Given $\tilde{x} \in \mathbb{R}^n$ and a function $F : \mathbb{R}_+ \to \mathbb{R}$, define the function $f : \mathbb{R}^n \to \mathbb{R}$ by $f(x) = F(\|x - \tilde{x}\|)$ for any $x \in \mathbb{R}^n$. Then, for any $\rho_2 \geq \rho_1 \geq 0$, one has*

$$\int_{\mathbf{B}_{\rho_2}(\tilde{x}) \setminus \mathbf{B}_{\rho_1}(\tilde{x})} f(x)dx = n\gamma_n \int_{\rho_1}^{\rho_2} z^{n-1} F(z)dz,$$

*where $\gamma_n = \frac{\pi^{(n-1)/2} 2^{(n+1)/2}}{n!!}$ is the volume of the unit Euclidean ball in $\mathbb{R}^n$.*

**Proof.** Apply a change of variables using spherical coordinates as explained, e.g., in [7]. ☐

**Lemma 5.12.** *For any positive integers $r$ and $n$, one has $\left(\frac{1}{2r+1}\right)^{-\frac{n}{4(2r+1)+2n}} < 6n$.*

**Proof.** Let $n \in \mathbb{N}$ be given. Denote

$$g(r) := \left(\frac{1}{2r+1}\right)^{-\frac{n}{4(2r+1)+2n}} = (2r+1)^{\frac{n}{4(2r+1)+2n}} \quad (r \geq 0).$$

Observe that, $g(0) = 1$, $g(r) > 0$ for all $r \geq 0$, $\ln(g(r)) = \frac{n}{8r+4+2n} \ln(2r+1)$, and thus $\lim_{r\to\infty} g(r) = 1$. It suffices to show $g(r^*) < 6n$ for any stationary point $r^*$. Since

$$\frac{d \ln(g(r))}{dr} = \frac{-8n \ln(2r+1)}{(8r+4+2n)^2} + \frac{2n}{(2r+1)(8r+4+2n)},$$

and $g'(r) = \frac{1}{g(r)} \frac{d \ln(g(r))}{dr}$, any stationary point $r^*$ satisfies

$$\frac{d \ln(g(r^*))}{dr} = 0 \Longleftrightarrow (2r^*+1)\left[\ln(2r^*+1) - 1\right] = \frac{n}{2}.$$

Since

$$(2r^*+1)(\ln(3) - 1) \leq (2r^*+1)\left[\ln(2r^*+1) - 1\right] = \frac{n}{2},$$

one has $2r^* + 1 \leq \frac{n}{2(\ln(3)-1)} < 6n$. Since $g(r) \leq 2r+1$ for all $r \geq 0$, one has $g(r^*) \leq 2r^* + 1 < 6n$. ☐

**Lemma 5.13.** *Assume $\mathbf{K} \subseteq \mathbb{R}^n$ is compact and satisfies Assumption 5.4. Then, for any $0 < \epsilon \leq \epsilon_{\mathbf{K}}$, one has*

$$\int_{\mathbf{K}} C_{\mathbf{K},a} \|x - a\| G_a(x) dx \leq \epsilon + \frac{n\sigma^{n+1} p(n)}{\epsilon^n \eta_{\mathbf{K}}} e^{\frac{\epsilon^2}{2\sigma^2}},$$

*where $p(n) := \int_0^{+\infty} t^n e^{-t^2/2} dt$ is given by*

$$p(n) = \begin{cases} 1 & \text{if } n = 1, \\ \sqrt{\frac{\pi}{2}} \prod_{j=1}^{k} (2j-1) & \text{if } n = 2k \text{ and } k \geq 1, \\ \prod_{j=1}^{k} (2j) & \text{if } n = 2k+1 \text{ and } k \geq 1. \end{cases} \tag{5.24}$$

**Proof.** Let $\varphi := \int_{\mathbf{K}} C_{\mathbf{K},a} \|x - a\| G_a(x) dx$ denote the integral that we need to upper bound. We split the integral $\varphi$ as $\varphi = \varphi_1 + \varphi_2$, depending on whether $x$ lies in the ball $\mathbf{B}_\epsilon(a)$ or not.

First, we upper bound the term $\varphi_1$ as

$$\varphi_1 := \int_{\mathbf{K} \cap \mathbf{B}_\epsilon(a)} \|x - a\| C_{\mathbf{K},a} G_a(x) dx \leq \epsilon \int_{\mathbf{K} \cap \mathbf{B}_\epsilon(a)} C_{\mathbf{K},a} G_a(x) dx \leq \epsilon \int_{\mathbf{K}} C_{\mathbf{K},a} G_a(x) dx = \epsilon.$$

Second, we bound the integral

$$\varphi_2 := C_{\mathbf{K},a} \int_{\mathbf{K} \setminus \mathbf{B}_\epsilon(a)} \|x - a\| G_a(x) dx.$$

Since $\mathbf{K} \subseteq \mathbf{B}_{\sqrt{D(\mathbf{K})}}(a)$, one has

$$\varphi_2 \leq C_{\mathbf{K},a} \int_{\mathbf{B}_{\sqrt{D(\mathbf{K})}}(a) \setminus \mathbf{B}_\epsilon(a)} \|x - a\| G_a(x) dx,$$

where the right hand side, by Lemma 5.11, is equal to

$$\frac{C_{\mathbf{K},a} n \gamma_n}{(2\pi\sigma^2)^{n/2}} \int_\epsilon^{\sqrt{D(\mathbf{K})}} z^n \exp\left(-\frac{z^2}{2\sigma^2}\right) dz.$$

By a change of variable $t = \frac{z}{\sigma}$, one obtains

$$\varphi_2 \leq \frac{C_{\mathbf{K},a} n \gamma_n \sigma}{(2\pi)^{n/2}} \int_{\epsilon/\sigma}^{\sqrt{D(\mathbf{K})}/\sigma} t^n \exp\left(-\frac{t^2}{2}\right) dt,$$

and thus

$$\varphi_2 \leq \frac{C_{\mathbf{K},a} n \gamma_n \sigma}{(2\pi)^{n/2}} \int_0^{+\infty} t^n \exp\left(-\frac{t^2}{2}\right) dt = \frac{C_{\mathbf{K},a} n \gamma_n \sigma}{(2\pi)^{n/2}} p(n).$$

Here we have set $p(n) := \int_0^{+\infty} t^n e^{-\frac{t^2}{2}} dt$ which can be checked to be given by (5.24) (e.g., using induction on $n$). Now, combining with the upper bound for $C_{\mathbf{K},a}$ from (5.23), we obtain

$$\varphi_2 \leq \frac{n\sigma^{n+1} p(n)}{\epsilon^n \eta_{\mathbf{K}}} e^{\frac{\epsilon^2}{2\sigma^2}}.$$

Therefore, we have shown:

$$\varphi = \varphi_1 + \varphi_2 \leq \epsilon + \frac{n\sigma^{n+1} p(n)}{\epsilon^n \eta_{\mathbf{K}}} e^{\frac{\epsilon^2}{2\sigma^2}},$$

which shows the lemma. $\square$

We are now ready to prove Theorem 5.9.

**Proof.** *(of Theorem 5.9)* Observe that, if $f$ is a polynomial, then we can use the upper bound (5.10) for its Lipschitz constant and thus the inequality (5.22) follows as a direct consequence of the inequality (5.21). Therefore, it suffices to show the relation (5.21).

Recall that $a$ is a minimizer of $f$ over $\mathbf{K}$. As $f$ is Lipschitz continuous with Lipschitz constant $M_f$ on $\mathbf{K}$, we have

$$f(x) - f(a) \leq M_f \|x - a\| \quad \forall x \in \mathbf{K}.$$

This implies

$$f^{(r)}_{\mathbf{K},a} - f_{\min,\mathbf{K}} = \int_{\mathbf{K}} c^r_{\mathbf{K},a} H_{r,a}(x)(f(x) - f(a)) dx \leq M_f \int_{\mathbf{K}} \|x - a\| c^r_{\mathbf{K},a} H_{r,a}(x) dx.$$

Our objective is now to show the existence of a constant $\zeta(\mathbf{K})$ such that

$$\psi := \int_{\mathbf{K}} c^r_{\mathbf{K},a} \|x - a\| H_{r,a}(x) dx \leq \frac{\zeta(\mathbf{K})}{\sqrt{2r+1}}, \quad \text{for any } r \geq r_{\mathbf{K}}$$

(recall that $r_{\mathbf{K}}$ is defined in (5.12)), by which we can then conclude the proof for (5.21).

For this, we split the integral $\psi$ as the sum of two terms:

$$\psi = \underbrace{\int_{\mathbf{K}} c^r_{\mathbf{K},a}\|x-a\|G_a(x)dx}_{=:\psi_1} + \underbrace{\int_{\mathbf{K}} c^r_{\mathbf{K},a}\|x-a\|(H_{r,a}(x)-G_a(x))dx}_{=:\psi_2}.$$

First, we upper bound the term $\psi_1$. As $c^r_{\mathbf{K},a} \leq C_{\mathbf{K},a}$ (by (5.23)), we can use Lemma 5.13 to conclude that, for any $0 < \epsilon \leq \epsilon_{\mathbf{K}}$,

$$\psi_1 \leq \int_{\mathbf{K}} C_{\mathbf{K},a}\|x-a\|G_a(x)dx \leq \epsilon + \frac{n\sigma^{n+1}p(n)}{\epsilon^n \eta_{\mathbf{K}}} e^{\frac{\epsilon^2}{2\sigma^2}} = \epsilon \underbrace{\left[1 + \frac{n\sigma^{n+1}p(n)}{\epsilon^{n+1}\eta_{\mathbf{K}}} e^{\frac{\epsilon^2}{2\sigma^2}}\right]}_{=:\mu_1} = \epsilon\mu_1. \tag{5.25}$$

Second we bound the integral

$$\psi_2 = \int_{\mathbf{K}} c^r_{\mathbf{K},a}\|x-a\|(H_{r,a}(x)-G_a(x))dx.$$

We can upper bound the function $H_{r,a}(x) - G_a(x)$ using the estimate from (5.19) and we get

$$H_{r,a}(x) - G_a(x) \leq \frac{1}{(2\pi\sigma^2)^{n/2}} \frac{\|x-a\|^{4r+2}}{(2\sigma^2)^{2r+1}(2r+1)!}.$$

Then we have

$$\begin{aligned} \psi_2 &\leq \frac{1}{(2\pi\sigma^2)^{n/2}} \int_{\mathbf{K}} c^r_{\mathbf{K},a} \frac{\|x-a\|^{4r+3}}{(2\sigma^2)^{2r+1}(2r+1)!} dx \\ &= \frac{1}{(2\pi\sigma^2)^{n/2}} \frac{c^r_{\mathbf{K},a}}{(2\sigma^2)^{2r+1}(2r+1)!} \int_{\mathbf{K}} \|x-a\|^{4r+3} dx. \end{aligned}$$

Now we upper bound the integral $\int_{\mathbf{K}} \|x-a\|^{4r+3}dx$. Since $\mathbf{K} \subseteq \mathbf{B}_{\sqrt{D(\mathbf{K})}}(a)$, one has

$$\int_{\mathbf{K}} \|x-a\|^{4r+3}dx \leq \int_{\mathbf{B}_{\sqrt{D(\mathbf{K})}}(a)} \|x-a\|^{4r+3}dx,$$

where the right hand side, by Lemma 5.11, is equal to

$$n\gamma_n \int_0^{\sqrt{D(\mathbf{K})}} z^{4r+n+2}dz = \frac{n\gamma_n D(\mathbf{K})^{\frac{4r+n+3}{2}}}{4r+n+3} \leq n\gamma_n D(\mathbf{K})^{\frac{4r+n+3}{2}}.$$

Thus, we obtain

$$\psi_2 \leq \frac{1}{(2\pi\sigma^2)^{n/2}} \frac{c^r_{\mathbf{K},a}}{(2\sigma^2)^{2r+1}(2r+1)!} n\gamma_n D(\mathbf{K})^{\frac{4r+n+3}{2}}.$$

We now use the upper bound for $c^r_{\mathbf{K},a}$ from (5.23):

$$c^r_{\mathbf{K},a} \leq \frac{(2\pi\sigma^2)^{n/2}\exp\left(\frac{\epsilon^2}{2\sigma^2}\right)}{\eta_{\mathbf{K}}\epsilon^n\gamma_n}$$

and we obtain

$$\psi_2 \leq \frac{n\exp\left(\frac{\epsilon^2}{2\sigma^2}\right) D(\mathbf{K})^{\frac{4r+n+3}{2}}}{\eta_{\mathbf{K}}\epsilon^n(2r+1)!(2\sigma^2)^{2r+1}}.$$

Finally we use the Stirling's inequality:

$$(2r+1)! \geq \sqrt{2\pi(2r+1)}\left(\frac{2r+1}{e}\right)^{2r+1},$$

and obtain

$$\begin{aligned} \psi_2 &\leq \underbrace{\frac{n\exp\left(\frac{\epsilon^2}{2\sigma^2}\right) D(\mathbf{K})^{\frac{n+1}{2}}}{\eta_{\mathbf{K}}}}_{=:\mu_2}\left(\frac{D(\mathbf{K})e}{2\sigma^2\epsilon^{n/(2r+1)}(2r+1)}\right)^{2r+1}\frac{1}{\sqrt{2\pi(2r+1)}} \qquad (5.26)\\ &= \frac{\mu_2}{\sqrt{2\pi(2r+1)}}\left(\frac{D(\mathbf{K})e}{2\sigma^2\epsilon^{n/(2r+1)}(2r+1)}\right)^{2r+1}. \end{aligned}$$

We can now upper bound the quantity $\psi = \psi_1 + \psi_2$, by combining the upper bound for $\psi_1$ in (5.25) with the above upper bound (5.26) for $\psi_2$. That is,

$$\psi \leq \epsilon\mu_1 + \frac{\mu_2}{\sqrt{2\pi(2r+1)}}\left(\frac{D(\mathbf{K})e}{2\sigma^2\epsilon^{n/(2r+1)}(2r+1)}\right)^{2r+1}.$$

We now indicate how to select the parameters $\epsilon$ and $\sigma$.
First we select $\sigma = \epsilon$, so that both parameters $\mu_1$ and $\mu_2$ appearing in (5.25) and (5.26) are constants depending on $n$ and $\mathbf{K}$, namely

$$\mu_1 = 1 + \frac{np(n)e^{1/2}}{\eta_{\mathbf{K}}} \quad \text{and} \quad \mu_2 = \frac{ne^{1/2}D(\mathbf{K})^{\frac{n+1}{2}}}{\eta_{\mathbf{K}}}.$$

Next we select $\epsilon$ so that $\frac{D(\mathbf{K})e}{2\epsilon^{2+n/(2r+1)}(2r+1)} = 1$, i.e.,

$$\epsilon = \left(\frac{D(\mathbf{K})e}{2(2r+1)}\right)^{\frac{2r+1}{2(2r+1)+n}} = \left(\frac{D(\mathbf{K})e}{2}\right)^{\frac{2r+1}{2(2r+1)+n}} \left(\frac{1}{2r+1}\right)^{\frac{1}{2}-\frac{n}{4(2r+1)+2n}}.$$

Summarizing, we have shown that

$$\begin{aligned} \psi &\leq \left(\frac{1}{2r+1}\right)^{\frac{1}{2}-\frac{n}{4(2r+1)+2n}} \left[\left(\frac{D(\mathbf{K})e}{2}\right)^{\frac{2r+1}{2(2r+1)+n}} \mu_1 + \frac{\mu_2}{\sqrt{2\pi}}\left(\frac{1}{2r+1}\right)^{\frac{n}{4(2r+1)+2n}}\right] \\ &\leq \left(\frac{1}{2r+1}\right)^{\frac{1}{2}} 6n\left(\mu_1 \max\left\{1, \sqrt{\frac{D(\mathbf{K})e}{2}}\right\} + \frac{\mu_2}{\sqrt{2\pi}}\right). \end{aligned} \tag{5.27}$$

To obtain the last inequality (5.27), we use the inequality $\left(\frac{1}{2r+1}\right)^{-\frac{n}{4(2r+1)+2n}} < 6n$ (recall Lemma 5.12), together with the two inequalities

$$\left(\frac{D(\mathbf{K})e}{2}\right)^{\frac{2r+1}{2(2r+1)+n}} \leq \max\left\{1, \sqrt{\frac{D(\mathbf{K})e}{2}}\right\} \text{ and } \left(\frac{1}{2r+1}\right)^{\frac{n}{4(2r+1)+2n}} \leq 1.$$

Since we have assumed $\epsilon \leq \epsilon_{\mathbf{K}}$ (recall Lemma 5.10), this implies the condition

$$r \geq \frac{D(\mathbf{K})e}{4}\epsilon_{\mathbf{K}}^{-\left(2+\frac{n}{2r+1}\right)} - \frac{1}{2},$$

i.e., the inequality (5.27) holds for any $r \geq \frac{D(\mathbf{K})e}{4}\epsilon_{\mathbf{K}}^{-\left(2+\frac{n}{2r+1}\right)} - \frac{1}{2}$. If $\epsilon_{\mathbf{K}} \leq 1$ and $r \geq n/2$, then we have $\epsilon_{\mathbf{K}}^{-(2+\frac{n}{2r+1})} \leq \epsilon_{\mathbf{K}}^{-3}$ and thus the inequality (5.27) holds for any $r \geq \max\left\{\frac{D(\mathbf{K})e}{4\epsilon_{\mathbf{K}}^3}, \frac{n}{2}\right\}$. If $\epsilon_{\mathbf{K}} \geq 1$ then $\epsilon_{\mathbf{K}}^{-(2+\frac{n}{2r+1})} \leq 1$ and thus (5.27) holds for any integer $r \geq \frac{D(\mathbf{K})e}{4}$. Hence, the inequality (5.27) holds for any $r \geq r_{\mathbf{K}}/2$, where $r_{\mathbf{K}}$ is as defined in (5.12).
Finally, by defining the constant

$$\zeta(\mathbf{K}) := 6n\left(\mu_1 \max\left\{1, \sqrt{\frac{D(\mathbf{K})e}{2}}\right\} + \frac{\mu_2}{\sqrt{2\pi}}\right),$$

which indeed depends only on $\mathbf{K}$ and its dimension $n$, we can conclude the proof for (5.21). □

**Remark 5.14.** *Our main result in Theorem 5.7 shows the bound*

$$\underline{f}_{\mathbf{K}}^{(r)} - f_{\min,\mathbf{K}} = O\left(\frac{1}{\sqrt{r}}\right).$$

*For this we have used the sum of squares density $H_{r,a}(x)$ from (5.17), where $a$ is a global minimizer of $f$ over $\mathbf{K}$. A natural question is whether another choice of the density function might lead to a better bound for the convergence rate. We have in fact also considered the following density function*

$$\tilde{H}_{r,a}(x) = \left(1 - \frac{\|x-a\|^2}{D(\mathbf{K})}\right)^{2r},$$

*which is a sum of squares of degree $4r$. A motivation for trying this function is that it converges to 0 as $r \to \infty$ at all points $x \in \mathbf{K}\backslash\{a\}$. However, using this function $\tilde{H}_{r,a}(x)$, we could only prove a weaker bound, in the order $O\left(1/r^{\frac{1}{n+1}}\right)$, for the convergence rate of the parameters $\underline{f}_{\mathbf{K}}^{(r)}$.*

**Remark 5.15.** *Note that in the proof of Theorem 5.9, we use Assumption 5.4 only for the selected minimizer $a \in \mathbf{K}$ (and we use it only in the proof of Lemma 5.10). Hence, if the selected point $a$ lies in the interior of $\mathbf{K}$, i.e., if there exists $\delta > 0$ such that $\mathbf{B}_\delta(a) \subseteq \mathbf{K}$, then the result of Theorem 5.9 (and thus Theorem 5.7) holds when selecting $\eta_{\mathbf{K}} = 1$ and $\epsilon_{\mathbf{K}} = \delta$.*

*Our results extend also to unconstrained global minimization:*

$$f_{\min,\mathbb{R}^n} = \min_{x\in\mathbb{R}^n} f(x),$$

*if we know that $f$ has a global minimizer $a$ and we know a ball containing $a$. We can then indeed minimize $f$ over a compact set $\mathbf{K}$, which can be chosen to be the ball or a suitable hypercube containing $a$.*

## 5.3 Revisiting the main assumption

In this section we consider in more detail our Assumption 5.4. First we recall another condition, known as the *interior cone condition*, which is classically used in approximation theory (see, e.g., Wendland [97]).

**Definition 5.16.** *[97, Definition 3.1] A set $\mathbf{K} \subseteq \mathbb{R}^n$ is said to satisfy an interior cone condition if there exist an angle $\theta \in (0, \pi/2)$ and a radius $\rho > 0$ such that, for every $x \in \mathbf{K}$, a unit vector $\xi(x)$ exists such that the set*

$$C(x, \xi(x), \theta, \rho) := \{x + \lambda y : y \in \mathbb{R}^n, \|y\| = 1, y^T \xi(x) \geq \cos\theta, \lambda \in [0, \rho]\} \qquad (5.28)$$

*is contained in $\mathbf{K}$.*

For instance, every Euclidean ball satisfies the interior cone condition [97, Lemma 3.10].

**Lemma 5.17.** *[97, Lemma 3.10] Every Euclidean ball with radius $r > 0$ satisfies an interior cone condition with radius $\rho = r$ and angle $\theta = \pi/3$.*

**Proof.** For completeness, we review the proof of Lemma 5.17 given in [97].

We can assume w.l.o.g. that the ball is centered at zero. For every point $x$ in the ball we have to find a cone with prescribed radius and angle. For the center $x = 0$ we can choose any direction to see that such a cone is indeed contained in the ball. For $x \neq 0$ we choose the direction $\xi(x) = -x/\|x\|$, see Figure 5.3. A typical point on the cone is given by $x + \lambda y$ with $\|y\| = 1$, $y^T \xi(x) \geq \cos(\pi/3) = 1/2$ and $0 \leq \lambda \leq r$. For this point we find

$$\|x + \lambda y\|^2 = \|x\|^2 + \lambda^2 - 2\lambda\|x\|\xi(x)^T y \leq \|x\|^2 + \lambda^2 - \lambda\|x\|.$$

The last expression equals $\|x\|(\|x\| - \lambda) + \lambda^2$, which can be bounded by $\lambda^2 \leq r^2$ in the case $\|x\| \leq \lambda$. If $\|x\| \geq \lambda$ then we can transform the last expression to $\lambda(\lambda - \|x\|) + \|x\|^2$, which can be bounded by $\|x\|^2 \leq r^2$. Thus $x + \lambda y$ is contained in the ball. □

Moreover, one can show that any *star-shaped* set satisfies the interior cone condition, see, e.g., [97, Proposition 11.26].

**Definition 5.18.** *[97, Definition 11.25] A set $\mathbf{K}$ is said to be star-shaped with respect to a ball $\mathbf{B}_r(x_c)$ if, for every $x \in \mathbf{K}$, the closed convex hull of $\{x\} \cup \mathbf{B}_r(x_c)$ is contained in $\mathbf{K}$.*

**Proposition 5.19.** *[97, Proposition 11.26] If $\mathbf{K}$ is bounded, star-shaped with respect to a ball $\mathbf{B}_r(x_c)$, then $\mathbf{K}$ satisfies an interior cone condition with radius $\rho = r$ and angle $\theta = 2\arcsin\left[\frac{r}{2\sqrt{D(\mathbf{K})}}\right]$.*

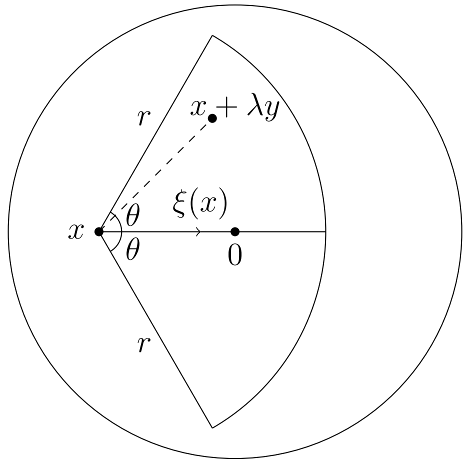


Figure 5.3: Euclidean Balls satisfy an interior cone condition

**Proof.** For completeness, we review the proof of Proposition 5.19 given in [97].
(i) When $x \in \mathbf{B}_r(x_c)$, it follows from Lemma 5.17 that $\mathbf{K}$ contains a cone pointed at $x$ with radius $r$ and angle $\pi/3$. It suffices now to observe that $\pi/3$ is at least the selected angle $\theta = 2 \arcsin \left[\frac{r}{2\sqrt{D(\mathbf{K})}}\right]$, since $r \leq \sqrt{D(\mathbf{K})}$.
(ii) If $x$ is outside the ball $\mathbf{B}_r(x_c)$, then we consider the convex hull of $x$ and the intersection of the sphere $S(x, \|x - x_c\|) = \{y \in \mathbb{R}^n : \|y - x\| = \|x_c - x\|\}$ with $\mathbf{B}_r(x_c)$, see Figure 5.4. This is a cone and, because $\mathbf{K}$ is star-shaped with respect to $\mathbf{B}_r(x_c)$, it is contained in $\mathbf{K}$. Its radius is the distance from $x$ to $x_c$. To find its angle $\theta$, we consider a triangle formed by $x, x_c$, and any point $y$ in the intersection of $S(x, \|x - x_c\|)$ and the sphere $S(x_c, r)$. This is an isosceles triangle, since $\|y - x\| = \|x_c - x\|$. The angle is $\theta = \angle x_c x y$; the side opposite this angle has length $r$. A little trigonometry then gives us $\|x_c - x\| \sin(\theta/2) = r/2$. Consequently, we have $\theta = 2 \arcsin \left[\frac{r}{2\|x_c - x\|}\right]$. Moreover, since $\|x_c - x\| \leq \sqrt{D(\mathbf{K})}$, we have $\theta \geq 2 \arcsin \left[\frac{r}{2\sqrt{D(\mathbf{K})}}\right]$.
This finishes the proof. $\square$

In fact, any set satisfying the interior cone condition also satisfies Assumption 5.4.

**Lemma 5.20.** *If a set $\mathbf{K} \subseteq \mathbb{R}^n$ satisfies the interior cone condition (5.28) then $\mathbf{K}$ also satisfies Assumption 5.4, where we set*

$$\eta_{\mathbf{K}} = \left[\frac{\sin \theta}{1 + \sin \theta}\right]^n \quad \textit{and} \quad \epsilon_{\mathbf{K}} = \rho.$$

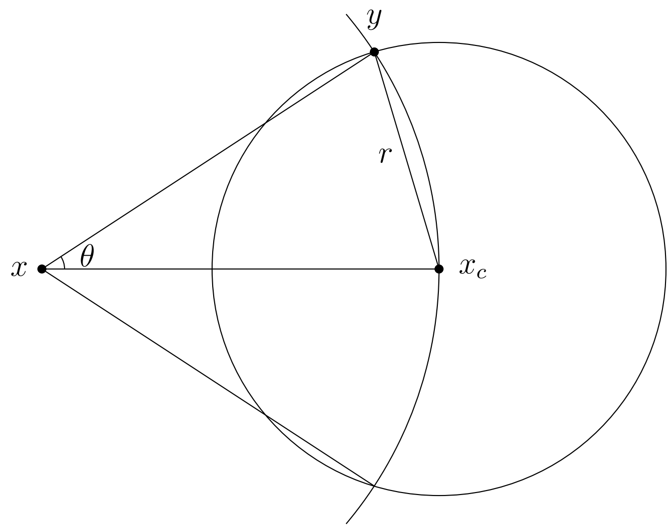


Figure 5.4: Bounded star-shaped sets satisfy the interior cone condition.

**Proof.** Assume that $\mathbf{K}$ satisfies the interior cone condition (5.28). Then, using [97, Lemma 3.7], we know that, for every $x \in \mathbf{K}$ and $h \leq \rho/(1+\sin\theta)$, the closed ball $\mathbf{B}_{h\sin\theta}(x + h\xi(x))$ is contained in $C(x, \xi(x), \theta, \rho)$ and thus in $\mathbf{K}$. Then, for any $x_0 \in \mathbf{K}$ and $\epsilon \in (0, \rho]$, after setting $h = \epsilon/(1+\sin\theta)$, one can obtain

$$\frac{\mathrm{vol}(\mathbf{B}_\epsilon(x_0) \cap \mathbf{K})}{\mathrm{vol}\mathbf{B}_\epsilon(x_0)} \geq \frac{\mathrm{vol}C(x_0, \xi(x_0), \theta, \epsilon)}{\mathrm{vol}\mathbf{B}_\epsilon(x_0)} \geq \frac{\mathrm{vol}\mathbf{B}_{h\sin\theta}(x_0 + h\xi(x_0))}{\mathrm{vol}\mathbf{B}_\epsilon(x_0)} = \left[\frac{\sin\theta}{1+\sin\theta}\right]^n.$$

Thus, Assumption 5.4 holds after setting $\eta_{\mathbf{K}} = \left[\frac{\sin\theta}{1+\sin\theta}\right]^n$ and $\epsilon_{\mathbf{K}} = \rho$. □

Note that any full-dimensional convex set that contains a ball $\mathbf{B}_r(x_c)$ is star-shaped with respect to $\mathbf{B}_r(x_c)$. Then, by Proposition 5.19 and Lemma 5.20, any full-dimensional bounded convex set satisfies the interior cone condition and thus Assumption 5.4.

**Corollary 5.21.** *Full-dimensional bounded convex sets satisfy the interior cone condition and Assumption 5.4.*

For example, the hypercube $\mathbf{Q}_n = [0,1]^n$ is a full-dimensional compact convex set and, by Proposition 5.19, it satisfies the interior cone condition with radius $\rho = 1/2$ and angle $\theta = 2\arcsin\left[\frac{1}{4\sqrt{n}}\right]$. Similarly, one can also check that the full-dimensional simplex $\widehat{\Delta_n}$ satisfies the interior cone condition with radius $\rho = \frac{1}{n+\sqrt{n}}$ and angle $\theta = 2\arcsin\left[\frac{1}{2\sqrt{2}(n+\sqrt{n})}\right]$.

# 5.4 Sampling feasible solutions

In this section we indicate how to sample feasible points in the set $\mathbf{K}$ from the optimal density function obtained by solving the semidefinite program (5.4).

Let $f \in \mathbb{R}[x]$ be a polynomial. Suppose $h^*(x) \in \Sigma[x]_r$ is an optimal solution of the program (5.4), i.e., $\underline{f}_{\mathbf{K}}^{(r)} = \int_{\mathbf{K}} f(x)h^*(x)dx$ and $\int_{\mathbf{K}} h^*(x)dx = 1$. Then $h^*$ can be seen as the probability density function of a probability distribution on $\mathbf{K}$, denoted as $\mathcal{T}_{\mathbf{K}}$ and, for any random vector $X = (X_1, \dots, X_n) \sim \mathcal{T}_{\mathbf{K}}$, the expectation of $f(X)$ is given by:

$$\mathbb{E}\left[f(X)\right] = \int_{\mathbf{K}} f(x)h^*(x)dx = \underline{f}_{\mathbf{K}}^{(r)}. \tag{5.29}$$

In what follows, we recall a well-known method to generate random samples $x \in \mathbf{K}$ from the distribution $\mathcal{T}_{\mathbf{K}}$ – namely, the *method of conditional distributions* (see, e.g., [60, Section 8.5.1]). Then we will observe that with high probability one of these sample points satisfies (roughly) the inequality $f(x) \leq \underline{f}_{\mathbf{K}}^{(r)}$ (see Theorem 5.24 for details).

We now describe how to use the method of conditional distributions to sample a random vector $X = (X_1, \dots, X_n) \sim \mathcal{T}_{\mathbf{K}}$. Assume that, for each $i = 2, \dots, n$, we know the cumulative conditional distribution of $X_i$ given that $X_j = x_j$ for $j = 1, \dots, i-1$, defined in terms of probabilities as

$$F_i(x_i \mid x_1, \dots, x_{i-1}) := \mathbf{Pr}\left[X_i \leq x_i \;\mid\; X_1 = x_1, \dots, X_{i-1} = x_{i-1}\right].$$

Additionally, we assume that we know the cumulative marginal distribution function of $X_i$, defined as:

$$F_i(x_i) := \mathbf{Pr}\left[X_i \leq x_i\right].$$

Then one can generate a random sample $x = (x_1, \dots, x_n) \in \mathbf{K}$ from the distribution $\mathcal{T}_{\mathbf{K}}$ by the following algorithm:

- Generate $x_1$ with cumulative distribution function $F_1(\cdot)$.
- Generate $x_2$ with cumulative distribution function $F_2\left(\cdot|x_1\right)$.

  $\vdots$

- Generate $x_n$ with cumulative distribution function $F_n\left(\cdot|x_1, \dots, x_{n-1}\right)$.

Then return $x = (x_1, x_2, \ldots, x_n)^T$.

There remains to explain how to generate a (univariate) sample point $x$ with a given cumulative distribution function $F(\cdot)$, since this operation is carried out at each of the $n$ steps of the above algorithm. For this one can use the classical *inverse-transform method* (see, e.g., [60, Section 8.2.1]), which can be described as follows:

- Generate a sample $u$ from the uniform distribution over $[0, 1]$.
- Return $x = F^{-1}(u)$ if $F$ is strictly monotone increasing, or $x = \min\{y : F(y) \geq u\}$ otherwise.

To sample from the uniform distribution on $[0, 1]$, one can use *pseudo-random generators* in practice (see, e.g., [60, Chapter 7]). Moreover, in our setting, the function $F(\cdot)$ is a univariate polynomial (see Example 5.23 below). Thus solving the univariate equation $x = F^{-1}(u)$ reduces to computing the eigenvalues of the corresponding companion matrix (see, e.g., [58, Section 2.4.1]).

**Remark 5.22.** *In the formal setting of a randomized Turing machine (where unbiased coin flips are allowed as operations), one may generate the decimal expansion of a random $x$ from the uniform distribution on $[0, 1]$ via an infinite series of coin flips. If this procedure is terminated after $n$ steps, then one is in fact selecting an interval of length $2^{-n}$ uniformly at random from the discrete set of the corresponding multi-partition of $[0, 1]$ into $2^n$ intervals.*

As an illustration, we now indicate how to compute the cumulative marginal and conditional distributions $F_i(\cdot)$ and $F_i(\cdot \mid x_1 \ldots x_{i-1})$ for the case of the hypercube $\mathbf{Q}_n = [0, 1]^n$. We will then apply this method to several examples of polynomial minimization over the hypercube in Section 5.5.

**Example 5.23.** *Suppose that we are given a sum of squares density function $h^*(x)$ on $\mathbf{Q}_n = [0, 1]^n$. For $i = 1, \ldots, n$, define the function $f_{1\ldots i} \in \mathbb{R}[x_1, \ldots, x_i]$ by*

$$f_{1\ldots i}(x_1, \ldots, x_i) = \int_0^1 \cdots \int_0^1 h^*(x_1, \ldots, x_n) dx_{i+1} \cdots dx_n.$$

*Then the cumulative marginal distribution function $F_1(\cdot)$ is given by*

$$F_1(x_1) = \int_0^{x_1} f_1(y) dy$$

*and, for $i = 2, \dots, n$, the cumulative conditional distribution function $F_i(\cdot \mid x_1 \dots x_{i-1})$ is given by*

$$F_i(x_i \mid x_1 \dots x_{i-1}) = \frac{\int_0^{x_i} f_{1\dots i}(x_1, \dots, x_{i-1}, y)dy}{f_{1\dots(i-1)}(x_1, \dots, x_{i-1})}.$$

*For the case of the hypercube $\mathbf{Q}_n = [0,1]^n$, the cumulative marginal and conditional distributions $F_1(x_1)$ and $F_i(x_i \mid x_1 \dots x_{i-1})$ (for $i = 2, \dots, n$) are polynomials that can be computed in closed form (since we know explicitly the moments of the Lebesgue measure on the hypercube, recall* (5.7)*).*

We now observe that if we generate sufficiently many samples from the distribution $\mathcal{T}_{\mathbf{K}}$ then, with high probability, one of these samples is a point $x \in \mathbf{K}$ satisfying (roughly) $f(x) \le \underline{f}_{\mathbf{K}}^{(r)}$.

**Theorem 5.24.** *Let $X \sim \mathcal{T}_{\mathbf{K}}$. For any $\epsilon > 0$,*

$$\mathbf{Pr}\left[f(X) > \underline{f}_{\mathbf{K}}^{(r)} + \epsilon\left(\underline{f}_{\mathbf{K}}^{(r)} - f_{\min,\mathbf{K}}\right)\right] < \frac{1}{1+\epsilon}.$$

**Proof.** Let $X \sim \mathcal{T}_{\mathbf{K}}$ so that $\mathbb{E}\left[f(X)\right] = \underline{f}_{\mathbf{K}}^{(r)}$. Define the nonnegative random variable

$$Y := f(X) - f_{\min,\mathbf{K}}.$$

Then, one has $\mathbb{E}\left[Y\right] = \underline{f}_{\mathbf{K}}^{(r)} - f_{\min,\mathbf{K}}$. Given $\epsilon > 0$, the Markov Inequality (see, e.g., [69, Theorem 3.2]) implies

$$\mathbf{Pr}\left[Y \ge (1+\epsilon)\mathbb{E}\left[Y\right]\right] \le \frac{1}{1+\epsilon}.$$

This completes the proof. □

For given $\epsilon > 0$, if one samples $N$ times independently from $\mathcal{T}_{\mathbf{K}}$, one therefore obtains an $x \in \mathbf{K}$ such that

$$f(x) \le \underline{f}_{\mathbf{K}}^{(r)} + \epsilon\left(\underline{f}_{\mathbf{K}}^{(r)} - f_{\min,\mathbf{K}}\right)$$

with probability at least $1 - \left(\frac{1}{1+\epsilon}\right)^N$. For example, if $N \ge 1 + \frac{1}{\epsilon}$ then this probability is at least $1 - 1/e$.

## 5.5 Numerical examples

In this section, we consider several well-known polynomial test functions from global optimization that are listed in Table 5.1 [98].

| Name | Formula | Minimum ($f_{\min,\mathbf{K}}$) | Search domain ($\mathbf{K}$) |
|---|---|---|---|
| Booth Function | $f = (x_1 + 2x_2 - 7)^2 + (2x_1 + x_2 - 5)^2$ | $f(1,3) = 0$ | $[-10,10]^2$ |
| Matyas Function | $f = 0.26(x_1^2 + x_2^2) - 0.48x_1x_2$ | $f(0,0) = 0$ | $[-10,10]^2$ |
| Three–Hump Camel Function | $f = 2x_1^2 - 1.05x_1^4 + \frac{1}{6}x_1^6 + x_1x_2 + x_2^2$ | $f(0,0) = 0$ | $[-5,5]^2$ |
| Motzkin Polynomial | $f = x_1^4x_2^2 + x_1^2x_2^4 - 3x_1^2x_2^2 + 1$ | $f(\pm1,\pm1) = 0$ | $[-2,2]^2$ |
| Styblinski–Tang Function (with $n = 2,3,4$) | $f = \sum_{i=1}^n \frac{1}{2}x_i^4 - 8x_i^2 + \frac{5}{2}x_i$ | $f(-2.0935,\ldots,-2.0935) = -39.166n$ | $[-5,5]^n$ |
| Rosenbrock Function (with $n = 2,3,4$) | $f = \sum_{i=1}^{n-1} 100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2$ | $f(1,\ldots,1) = 0$ | $[-2.048, 2.048]^n$ |

Table 5.1: Test functions

For these functions, we calculate $\underline{f}_{\mathbf{K}}^{(r)}$ by solving the SDP (5.5) for increasing $r$.

We performed the computation on a PC with AMD Phenom(tm) 9600B Quad-Core CPU (2.30 GHz) and with 4 GB RAM. Moreover, we use CVX [35, 36] in MATLAB, selecting SDPT3 [93, 94] as the SDP solver.

We record the values $\underline{f}_{\mathbf{K}}^{(r)}$ as well as the CPU times (needed to solve the SDP) in Tables 5.2, 5.3 and 5.4.

Furthermore, for each order $r$, we use the method described in Section 5.4 to generate samples that are feasible solutions of (5.4), for the bivariate Rosenbrock and the Three–Hump Camel function in Table 5.1. For each order, the sample sizes 20 and 1000 are used. We also generate samples uniformly from the feasible set, for comparison. We give the results in Tables 5.5 and 5.6, where we record the mean, variance and the minimum value of these samples together with $\underline{f}_{\mathbf{K}}^{(r)}$ (which equals the sample mean by (5.29)).

Note that the average of the sample function values approximate $\underline{f}_{\mathbf{K}}^{(r)}$ reasonably

| r | Booth Function | | Matyas Function | | Three–Hump Camel Function | | Motzkin Polynomial | |
|---|---|---|---|---|---|---|---|---|
| | Value | Time (sec.) | Value | Time (sec.) | Value | Time (sec.) | Value | Time (sec.) |
| 1 | 244.680 | 0.30 | 8.26667 | 0.26 | 265.774 | 0.44 | 4.2 | 0.17 |
| 2 | 162.486 | 0.34 | 5.32223 | 0.34 | 29.0005 | 0.38 | 1.06147 | 0.28 |
| 3 | 118.383 | 0.41 | 4.28172 | 0.27 | 29.0005 | 0.31 | 1.06147 | 0.08 |
| 4 | 97.6473 | 0.39 | 3.89427 | 0.41 | 9.58064 | 0.39 | 0.829415 | 0.13 |
| 5 | 69.8174 | 0.55 | 3.68942 | 0.47 | 9.58064 | 0.55 | 0.801069 | 0.06 |
| 6 | 63.5454 | 0.59 | 2.99563 | 0.69 | 4.43983 | 0.55 | 0.801069 | 0.13 |
| 7 | 47.0467 | 0.64 | 2.54698 | 0.72 | 4.43983 | 0.59 | 0.708889 | 0.13 |
| 8 | 41.6727 | 0.70 | 2.04307 | 0.76 | 2.55032 | 0.67 | 0.565553 | 0.16 |
| 9 | 34.2140 | 0.83 | 1.83356 | 0.81 | 2.55032 | 0.70 | 0.565553 | 0.16 |
| 10 | 28.7248 | 0.94 | 1.47840 | 0.87 | 1.71275 | 0.84 | 0.507829 | 0.22 |
| 11 | 25.6050 | 1.03 | 1.37644 | 0.94 | 1.71275 | 0.84 | 0.406076 | 0.31 |
| 12 | 21.1869 | 1.48 | 1.11785 | 1.25 | 1.27749 | 1.11 | 0.406076 | 0.27 |

Table 5.2: $\underline{f}_{\mathbf{K}}^{(r)}$ for Booth, Matyas, Three-Hump Camel and Motzkin Functions

| r | Sty.–Tang ($n = 2$) | | Rosenb. ($n = 2$) | | Sty.–Tang ($n = 3$) | | Rosenb. ($n = 3$) | |
|---|---|---|---|---|---|---|---|---|
| | Value | Time (sec.) | Value | Time (sec.) | Value | Time (sec.) | Value | Time (sec.) |
| 1 | −12.9249 | 0.41 | 214.648 | 0.34 | −18.8832 | 0.34 | 629.086 | 0.37 |
| 2 | −25.7727 | 0.31 | 152.310 | 0.34 | −36.0339 | 0.38 | 394.187 | 0.34 |
| 3 | −34.4030 | 0.39 | 104.889 | 0.35 | −44.9525 | 0.65 | 295.811 | 0.44 |
| 4 | −41.4436 | 0.36 | 75.6010 | 0.33 | −54.4424 | 0.98 | 206.903 | 0.53 |
| 5 | −45.1032 | 0.41 | 51.5037 | 0.50 | −60.5823 | 0.66 | 168.135 | 0.66 |
| 6 | −51.0509 | 0.50 | 41.7878 | 0.45 | −67.6027 | 0.98 | 121.558 | 1.05 |
| 7 | −56.4050 | 0.52 | 30.1392 | 0.41 | −74.5791 | 1.33 | 101.953 | 1.23 |
| 8 | −58.6004 | 0.58 | 25.8329 | 0.42 | −79.1261 | 2.28 | 77.4797 | 1.92 |
| 9 | −60.7908 | 0.67 | 19.4972 | 0.55 | −82.9581 | 3.53 | 66.6954 | 3.08 |
| 10 | −64.0147 | 0.83 | 17.3999 | 0.61 | −87.6127 | 7.82 | 53.0369 | 4.44 |
| 11 | −65.7111 | 0.86 | 13.6289 | 0.76 | −91.0233 | 10.53 | 46.5871 | 7.89 |
| 12 | −66.5532 | 1.23 | 12.5024 | 0.94 | −93.2038 | 19.47 | 38.4281 | 13.99 |

Table 5.3: $\underline{f}_{\mathbf{K}}^{(r)}$ for Styblinski-Tang and Rosenbrock Functions (with $n = 2, 3$)

| r | Sty.–Tang ($n = 4$) | | Rosenb. ($n = 4$) | |
|---|---|---|---|---|
| | Value | Time (sec.) | Value | Time (sec.) |
| 1 | $-24.6541$ | 0.25 | 1048.19 | 0.34 |
| 2 | $-45.5192$ | 0.34 | 690.332 | 0.42 |
| 3 | $-55.0577$ | 0.61 | 536.367 | 0.48 |
| 4 | $-66.8202$ | 0.78 | 382.729 | 0.72 |
| 5 | $-74.7215$ | 1.37 | 314.758 | 1.39 |
| 6 | $-82.8699$ | 3.09 | 236.709 | 3.09 |
| 7 | $-90.8863$ | 9.98 | 202.674 | 6.61 |
| 8 | $-97.1192$ | 28.64 | 156.295 | 19.62 |
| 9 | $-102.387$ | 83.01 | 137.015 | 60.59 |

Table 5.4: $\underline{f}_{\mathbf{K}}^{(r)}$ for Styblinski-Tang and Rosenbrock Functions (with $n = 4$)

well for sample size 1000, but poorly for sample size 20. Moreover, the average sample function value for uniform sampling from $\mathbf{K}$ is much higher than $\underline{f}_{\mathbf{K}}^{(r)}$. Also, the minimum function value for sampling from $\mathcal{T}_{\mathbf{K}}$ is significantly lower than the minimum function value obtained by uniform sampling for most values of $r$. In terms of generating "good" feasible solutions, sampling from $\mathcal{T}_{\mathbf{K}}$ therefore outperforms uniform sampling from $\mathbf{K}$ for these examples, as one would expect.

## 5.6 Conclusion

In this chapter, we have shown $\underline{f}_{\mathbf{K}}^{(r)} - f_{\min,\mathbf{K}} = O\left(\frac{1}{\sqrt{r}}\right)$. However, we do not know any instance for which there is a lower bound for $\underline{f}_{\mathbf{K}}^{(r)} - f_{\min,\mathbf{K}}$ in the order $\Omega\left(\frac{1}{\sqrt{r}}\right)$. Thus it remains an open question to decide the exact rate of convergence for the parameter $\underline{f}_{\mathbf{K}}^{(r)}$, which is interesting for future research.

Recall that the computation of the upper bound $\underline{f}_{\mathbf{K}}^{(r)}$ by solving the semidefinite programs (5.5) involve matrix variables of order $\binom{n+r}{r}$. Thus one is limited to relatively small values of $n$ and $r$, when using interior point SDP solvers.

Having said that, the sampling approach of Section 5.4 often provides good feasible solutions for the examples in Section 5.5, even for small values of $r$. One may therefore explore using the sampling technique (for small $r$) as a way of generating starting points for multi-start global optimization algorithms.

Another possibility to enhance computation would be to investigate other sufficient

| r | $\underline{f}_{\mathbf{K}}^{(r)}$ | Mean | Variance | Minimum | Sample Size |
|---|---|---|---|---|---|
| 1 | 214.648 | 121.125 | 14005.5 | 0.00451826 | 20 |
| | | 209.9 | 80699.0 | 0.0008754 | 1000 |
| 2 | 152.310 | 184.496 | 58423.9 | 4.94265 | 20 |
| | | 149.6 | 54455.0 | 0.02805 | 1000 |
| 3 | 104.889 | 146.618 | 64611.2 | 0.0113339 | 20 |
| | | 110.1 | 26022.0 | 0.0665 | 1000 |
| 4 | 75.6010 | 62.4961 | 5803.21 | 0.0542813 | 20 |
| | | 75.65 | 45777.0 | 0.007285 | 1000 |
| 5 | 51.5037 | 58.4032 | 4397.0 | 0.668679 | 20 |
| | | 50.64 | 6285.0 | 0.01382 | 1000 |
| 6 | 41.7878 | 35.4183 | 2936.24 | 1.16154 | 20 |
| | | 37.64 | 3097.0 | 0.06188 | 1000 |
| 7 | 30.1392 | 29.6545 | 1022.2 | 1.05813 | 20 |
| | | 27.11 | 1332.0 | 0.02044 | 1000 |
| 8 | 25.8329 | 19.5392 | 301.334 | 0.505628 | 20 |
| | | 34.32 | 4106.0 | 0.074 | 1000 |
| 9 | 19.4972 | 20.8982 | 328.475 | 0.564992 | 20 |
| | | 18.65 | 593.6 | 0.07951 | 1000 |
| 10 | 17.3999 | 9.37959 | 146.496 | 0.562473 | 20 |
| | | 15.33 | 685.7 | 0.1448 | 1000 |
| 11 | 13.6289 | 8.74923 | 52.1436 | 0.75774 | 20 |
| | | 15.7 | 7498.0 | 0.1719 | 1000 |
| 12 | 12.5024 | 5.43151 | 66.561 | 0.438172 | 20 |
| | | 12.7 | 764.7 | 0.0945 | 1000 |
| Uniform Sample | | 489.722 | 433549.0 | 9.0754 | 20 |
| | | 465.729 | 361150.0 | 0.0771463 | 1000 |

Table 5.5: Sampling results for the Rosenbrock Function ($n = 2$)

conditions for nonnegativity of $h$ on $\mathbf{K}$, more general than the sum-of-squares condition studied here. For instance, if $\mathbf{K}$ is a semi-algebraic set defined as

$$\mathbf{K} = \{x \in \mathbb{R}^n : g_1(x) \geq 0, \ldots, g_m(x) \geq 0\},$$

then one may consider polynomials $h$ of the form $h = \sigma_0 + \sum_{i=1}^m g_i\sigma_i$ where $\sigma_i \in \Sigma[x]$ ($i \in \{0, 1, \ldots, m\}$). This may result in a faster rate of convergence than for $\underline{f}_{\mathbf{K}}^{(r)}$.

| r | $\underline{f}_{\mathbf{K}}^{(r)}$ | Mean | Variance | Minimum | Sample Size |
|---|---|---|---|---|---|
| 1 | 265.774 | 216.773 | 177142.0 | 0.106854 | 20 |
| | | 261.23 | 193466.0 | 0.11705 | 1000 |
| 2 | 29.0005 | 28.0344 | 2964.85 | 1.1718 | 20 |
| | | 27.712 | 6712.8 | 0.014255 | 1000 |
| 3 | 29.0005 | 14.9951 | 523.904 | 0.452655 | 20 |
| | | 32.363 | 16681.0 | 0.0088426 | 1000 |
| 4 | 9.58064 | 2.99756 | 14.1201 | 0.175016 | 20 |
| | | 10.364 | 1944.0 | 0.010013 | 1000 |
| 5 | 9.58064 | 4.41907 | 14.1358 | 0.419394 | 20 |
| | | 9.1658 | 643.88 | 0.0015924 | 1000 |
| 6 | 4.43983 | 7.98481 | 245.089 | 0.126147 | 20 |
| | | 4.5791 | 493.12 | 0.0035581 | 1000 |
| 7 | 4.43983 | 3.96711 | 20.3193 | 0.260331 | 20 |
| | | 3.7911 | 57.847 | 0.0076111 | 1000 |
| 8 | 2.55032 | 2.18925 | 3.87943 | 0.0310113 | 20 |
| | | 2.2302 | 8.3767 | 0.0028817 | 1000 |
| 9 | 2.55032 | 1.38102 | 2.27433 | 0.138641 | 20 |
| | | 3.2217 | 812.18 | 0.00014805 | 1000 |
| 10 | 1.71275 | 1.03179 | 0.992636 | 0.0645815 | 20 |
| | | 1.5069 | 3.9581 | 0.0014225 | 1000 |
| 11 | 1.71275 | 1.30757 | 1.90985 | 0.0320489 | 20 |
| | | 1.6379 | 7.2518 | 0.0021144 | 1000 |
| 12 | 1.27749 | 0.841194 | 0.914514 | 0.0369565 | 20 |
| | | 1.2105 | 2.3 | 0.0005154 | 1000 |
| Uniform Sample | | 304.032 | 163021.0 | 1.65885 | 20 |
| | | 243.216 | 183724.0 | 0.00975034 | 1000 |

Table 5.6: Sampling results for the Three-Hump Camel Function

# Part III

# An Application in Graph Theory

# Chapter 6

# Handelman's hierarchy for the maximum stable set problem

In this part we consider the maximum stable set problem, a well-known NP-hard problem in graph theory. We study a global optimization approach, based on reformulating the stability number $\alpha(G)$ as the maximum of a square-free (or multilinear) quadratic polynomial on the hypercube $[0,1]^n$ as we saw earlier in (1.20). Namely, given a graph $G = (V, E)$, the maximum cardinality $\alpha(G)$ of a stable set in $G$ can be computed via the following polynomial optimization problem:

$$\alpha(G) = \max_{x \in [0,1]^{|V|}} \sum_{i \in V} x_i - \sum_{ij \in E} x_i x_j \tag{6.1}$$

(see Proposition 6.9 for the proof). We focus on the maximum stable set problem, since it is fundamental in the sense that any polynomial optimization problem on the Boolean hypercube can be transformed into a maximum stable set problem using the so-called conflict graph [11]. Moreover, Cornaz and Jost [13] give a direct explicit reformulation for the graph coloring problem as an instance of maximum stable set problem.

Algebraic approaches for the maximum stable set problem have been long studied; see, e.g. the early work of Lovász [64] and the more recent work of De Loera et al. [26], where Hilbert's Nullstellensatz plays a central role to show the non-existence of a solution to a system of polynomial equations. Other algebraic approaches, based on finding conditions for expressing positivity of polynomials, permit to construct upper bounds for the stability number. Depending on the type of positivity certificates one finds linear or semidefinite programming bounds (cf. e.g. [34, 25, 49, 55, 81, 90]).

We investigate a hierarchy of LP-based upper bounds for the stability number $\alpha(G)$, which are based on applying Handelman's result in Theorem 1.3 to the formulation (6.1). This approach for the maximum stable set problem was initiated by Park and Hong [78] (also in [77] for the maximum cut problem) and we will extend several of their results.

While several other linear or semidefinite programming hierarchical relaxations exist, a main motivation for focusing on the relaxations of Handelman type is that they appear to be easier to analyze. Indeed, as mentioned in Section 1.2.1 (recall Theorem 1.5), explicit error bounds have been given for general polynomials in [18]. Then, as we will show in Section 6.1.2, sharper bounds that apply at any order of relaxation have been given in [77, 78] for square-free quadratic polynomials.

## 6.1 Introduction

In this chapter, we study the Handelman bound for the maximum stable set problem, which will be defined below in (6.12). In particular, we focus on the *rank* of the Handelman hierarchy, defined as the smallest order for which the Handelman bound is exact (see Definition 6.10 below). We bound the rank of the Handelman hierarchy for several graph classes, including perfect graphs, odd circuits and wheels, and their complements, in the general weighted case. Moreover we show that the Handelman bound of order 2 is equal to the fractional stability number (see Theorem 6.14). We also prove two different upper bounds for the Handelman rank for a weighted graph, one in terms of the (unweighted) stability number and one in terms of the weighted stability and fractional stability numbers (see Theorem 6.20 and Corollary 6.24).

In addition, for the maximum cut problem, we clarify how the Handelman hierarchy applies to the formulation (1.4) and show that it can be reformulated as optimization over a polytope defined by an explicit subset of valid inequalities for the cut polytope; as an application we find again several results of [77, 78] (see Section 6.4).

### 6.1.1 Square-free polynomial optimization over the hypercube

Now we group some observations about the Handelman hierarchy (introduced in Section 1.2.1) when it is applied to the problem of maximizing a square-free polynomial $f$ over the hypercube $\mathbf{Q}_n = [0,1]^n$:

$$f_{\max,\mathbf{Q}_n} = \max_{x\in[0,1]^n} f(x),$$

which can be reformulated as

$$f_{\max,\mathbf{Q}_n} = \min \lambda \text{ s.t. } \lambda - f \in \mathscr{P}(\mathbf{Q}_n),$$

where we recall that $\mathscr{P}(\mathbf{Q}_n)$ denotes the set of nonnegative polynomials over the set $\mathbf{Q}_n$.

In what follows we let $\mathcal{I}$ denote the ideal generated by the polynomials $x_i^2 - x_i$ for $i \in [n]$. Using the description of the hypercube by the inequalities: $x_i \geq 0, 1 - x_i \geq 0$ for $i \in [n]$, the corresponding Handelman set of order $r$ reads:

$$\mathcal{H}_r := \left\{ \sum_{\alpha,\beta \in \mathbb{N}^n : |\alpha+\beta| \leq r} c_{\alpha,\beta} x^\alpha (1-x)^\beta : c_{\alpha,\beta} \geq 0 \right\}. \tag{6.2}$$

We also consider the following subset consisting of all square-free polynomials in $\mathcal{H}_r$ involving only terms which do not lie in the ideal $\mathcal{I}$:

$$H_r := \left\{ \sum_{T \in \mathcal{P}_r(V), I \subseteq T} c_{T,I} x^I (1-x)^{T \setminus I} : c_{T,I} \geq 0 \right\}. \tag{6.3}$$

Clearly, in the definition of $H_r$, we can restrict without loss of generality to sets $T \in \mathcal{P}_{=r}(V)$. Indeed, if $T < r$, pick an element $k \in V \setminus T$ and elevate the degree of $x^I(1-x)^{T\setminus I}$ by writing $x^I(1-x)^{T\setminus I} = x^{I\cup\{k\}}(1-x)^{T\setminus I} + x^I(1-x)^{(T\setminus I)\cup\{k\}}$.

One can construct the following Handelman upper bound $\bar{f}^{(r)}_{\text{han}}$ for $f_{\max,\mathbf{Q}_n}$, using the set $\mathcal{H}_r$ in (6.2):

$$\bar{f}^{(r)}_{\text{han}} := \min\{\lambda : \lambda - f \in \mathcal{H}_r\}.$$

Indeed, for any polynomial in $\mathcal{H}_r$, it is nonnegative over $\mathbf{Q}_n$. Thus, for any integer $r \geq 1$, $\mathcal{H}_r \subseteq \mathscr{P}(\mathbf{Q}_n)$, and $\bar{f}^{(r)}_{\text{han}} \geq f_{\max,\mathbf{Q}_n}$.

We now show that $\bar{f}^{(r)}_{\text{han}}$ can alternatively be defined using the subset $H_r$ in (6.3).

**Proposition 6.1.** *Let $f \in \mathbb{R}[x]$ be a square-free polynomial. For any integer $r \geq 1$,*

$$\bar{f}^{(r)}_{\text{han}} = \min\{\lambda : \lambda - f \in H_r\}.$$

This result follows directly from Lemma 6.4 below, whose proof relies on the following Lemmas 6.2 and 6.3.

**Lemma 6.2.** *If $f$ is a square-free polynomial and $f \in \mathcal{I}$, then $f = 0$.*

**Proof.** We use induction on the number $n$ of variables. In the case $n = 1$, we have that $f = f_0 + f_1 x_1 = q_1 \cdot (x_1 - x_1^2)$, which implies $q_1 = 0$ and thus $f = 0$ by looking at the degrees of both sides. Suppose now that the result holds for $n = k-1$. Let $f$ be a square-free polynomial in $k$ variables lying in the ideal $\mathcal{I}$. We can write $f$ as $f(x) = f_0(\underline{x}) + x_k f_1(\underline{x})$, where $f_0$, $f_1$ are square-free in the $k-1$ variables $\underline{x} = (x_1, \cdots, x_{k-1})$. Say, $f_0 + x_k f_1 = f = \sum_{i=1}^{k} q_i \cdot (x_i - x_i^2)$ for some polynomials $q_i$. By setting $x_k = 0$ we get: $f_0(\underline{x}) = \sum_{i=1}^{k-1} q_i(\underline{x}, 0)(x_i - x_i^2)$. As $f_0$ is square-free, we deduce using the induction assumption that $f_0 = 0$. Next, by setting $x_k = 1$, we get: $f_1(\underline{x}) = \sum_{i=1}^{k-1} q_i(\underline{x}, 1)(x_i - x_i^2)$. As $f_1$ is square-free we deduce from the induction assumption that $f_1 = 0$. Thus we have shown that $f = 0$. □

**Lemma 6.3.** *Given $\alpha, \beta \in \mathbb{N}^n$, let $I = \{i \in [n] : \alpha_i \geq 1\}$ and $J = \{i \in [n] : \beta_i \geq 1\}$ denote their supports.*

*(i) If $I \cap J \neq \emptyset$ then $x^\alpha (1-x)^\beta$ belongs to $\mathcal{I}$.*

*(ii) If $I \cap J = \emptyset$ then $x^\alpha (1-x)^\beta - x^I (1-x)^J$ belongs to $\mathcal{I}$.*

**Proof.** (i) Say, $1 \in I \cap J$. Then $x_1(1-x_1)$ is a factor of $x^\alpha(1-x)^\beta$ and thus $x^\alpha (1-x)^\beta \in \mathcal{I}$.
(ii) The proof is based on using iteratively the following identities, for any $k \geq 2$:

$$x_i^k - x_i = (x_i^2 - x_i)(x_i^{k-2} + \cdots + x_i + 1) \in \mathcal{I},$$

$$(1-x_i)^k - (1-x_i) = -x_i(1-x_i)((1-x_i)^{k-2} + \cdots + (1-x_i) + 1) \in \mathcal{I}.$$

Indeed, $x^\alpha(1-x)^\beta - x^I(1-x)^J = (x_1^{\alpha_1} - x_1)\underline{x}^{\underline{\alpha}}(1-x)^\beta + x_1(\underline{x}^{\underline{\alpha}}(1-x)^\beta - \underline{x}^{I\setminus\{1\}}(1-x)^J)$, setting $\underline{x} = (x_2, \cdots, x_n)$ and $\underline{\alpha} = (\alpha_2, \cdots, \alpha_n)$. □

**Lemma 6.4.** *Let $f$ be a square-free polynomial and $r \geq 1$ an integer. The following assertions are equivalent.*

*(i) $f \in \mathcal{H}_r$.*

*(ii) $f \in H_r + \mathcal{I}$.*

*(iii) $f \in H_r$.*

**Proof.** (i) $\Longrightarrow$ (ii): Say, $f = \sum_A c_{\alpha,\beta} x^\alpha (1-x)^\beta$ where $c_{\alpha,\beta} \geq 0$. Group in the polynomial $f_0 = \sum_{A_0} c_{\alpha,\beta} x^\alpha (1-x)^\beta$ all the terms of $f$ where the supports of $\alpha$ and $\beta$ are not disjoint. Let $S_\alpha$ denote the support of $\alpha$. Then, we have:

$$f = f_0 + \sum_{A \setminus A_0} c_{\alpha,\beta}(x^\alpha (1-x)^\beta - x^{S_\alpha}(1-x)^{S_\beta}) + \sum_{A\setminus A_0} c_{\alpha,\beta} x^{S_\alpha}(1-x)^{S_\beta}.$$

By Lemma 6.3, the first two sums lie in $\mathcal{I}$ and the last sum lies in $H_r$ and thus $f \in H_r + \mathcal{I}$.
The implication (ii) $\Longrightarrow$ (iii) follows from Lemma 6.2 and (iii) $\Longrightarrow$ (i) follows from the inclusion $H_r \subseteq \mathcal{H}_r$. $\square$

As an application of Lemma 6.2, we also find the following representation for square-free polynomials, which corresponds to the fact that the polynomials $\{x^I (1-x)^{[n]\setminus I} : I \subseteq [n]\}$ form a basis of the vector space of square-free polynomials.

**Corollary 6.5.** *Any square-free polynomial $f$ can be written as*

$$f = \sum_{I \subseteq [n]} f(\chi^I) x^I (1-x)^{[n]\setminus I}. \tag{6.4}$$

*Therefore, if $f(x) \geq 0$ for all $x \in \{0,1\}^n$, then $f \in H_n$.*

**Proof.** The polynomial $f - \sum_{I\subseteq[n]} f(\chi^I) x^I (1-x)^{[n]\setminus I}$ is square-free and vanishes on $\{0,1\}^n$. Hence it belongs to the ideal $\mathcal{I}$ and thus it is identically zero, by Lemma 6.2. $\square$

In particular, as the polynomial $f_{\max,\mathbf{Q}_n} - f$ is nonnegative on the hypercube $[0,1]^n$, then by Corollary 6.5, we find the convergence $\bar{f}^{(n)}_{\text{han}} = f_{\max,\mathbf{Q}_n}$ in $n$ steps, when $f$ is square-free. We mention another application which we will use later in this chapter.

**Lemma 6.6.** *Let $f$ be a square-free polynomial in $n$ variables $x = (x_1, \ldots, x_n) = (\underline{x}, x_n)$, setting $\underline{x} = (x_1, x_2 \ldots, x_{n-1})$. Then, one has*

$$f(x) = (1-x_n) f(\underline{x}, 0) + x_n f(\underline{x}, 1).$$

**Proof.** Using (6.4) (and splitting the sum into two sums depending whether $I$ contains $n$ or not), we can write $f(x)$ as $f(x) = x_n f_1(\underline{x}) + (1-x_n) f_2(\underline{x})$. By evaluating $f$ at $(\underline{x}, 0)$ and $(\underline{x}, 1)$, we obtain that $f(\underline{x}, 0) = f_2(\underline{x})$ and $f(\underline{x}, 1) = f_1(\underline{x})$, which gives the result. $\square$

### 6.1.2 Error bound of Handelman hierarchy

As we saw earlier in Theorem 1.6, De Klerk and Laurent [18] show that for any quadratic polynomial $f = x^T Ax + b^T x$, one has

$$\bar{f}^{(tn)}_{\mathrm{han}} - f_{\max,\mathbf{Q}_n} \leq \frac{-\sum_{i:A_{ii}<0} A_{ii}}{t}, \text{ for any } t \geq 1.$$

Note that the above result holds only for relaxations of order $r \geq n$. Moreover, if $f$ is square-free quadratic polynomial (i.e., $A_{ii} = 0$ for all $i$), then we find again the convergence $\bar{f}^{(n)}_{\mathrm{han}} = f_{\max,\mathbf{Q}_n}$ in $n$ steps.
Using a combinatorial version of Bernstein approximations, Park and Hong [78] can analyze the Handelman bound of any order $r \leq n$, in the quadratic square-free case. More precisely, they show the following result.

**Theorem 6.7.** *[78] Let $f = x^T Ax + b^T x$ be a quadratic polynomial which is square-free, i.e., $A_{ii} = 0$ for all $i \in [n]$. Assume moreover that $A_{ij} \leq 0$ for all $i \neq j \in [n]$. Then, for any integer $2 \leq r \leq n$,*

$$\bar{f}^{(r)}_{\mathrm{han}} \leq \frac{n}{r} f_{\max,\mathbf{Q}_n}.$$

We now extend the result of Theorem 6.7 analyzing the Handelman bound of any order $r \leq n$ to polynomials of arbitrary degree.

**Theorem 6.8.** *Let $f = \sum_{J\subseteq[n]} f_J x^J$ be a square-free polynomial with $f(0) = 0$. For any integer $r$ satisfying $\deg(f) \leq r \leq n$, we have*

$$\bar{f}^{(r)}_{\mathrm{han}} \leq \frac{n}{r} f_{\max,\mathbf{Q}_n} + \sum_{J\subseteq[n]:|J|\geq 2, f_J>0} f_J \lambda_J,$$

*setting*

$$\lambda_J = \left( \binom{n-1}{r-1} - \binom{n-|J|}{r-|J|} \right) \Big/ \binom{n-1}{r-1} \quad \text{for } J \subseteq [n].$$

*Hence, if $f_J \leq 0$ for all $J \subseteq [n]$ with $|J| \geq 2$, then*

$$\bar{f}^{(r)}_{\mathrm{han}} \leq \frac{n}{r} f_{\max,\mathbf{Q}_n}.$$

**Proof.** The proof is along the same lines as the proof of [78, Proposition 3.2] and uses the following 'combinatorial' Bernstein approximation of $f$, defined as

$$B_r(f) := \sum_{T\in\mathcal{P}_{=r}([n])} \sum_{I\subseteq T} f(\chi^I) x^I (1-x)^{T\setminus I}.$$

One can check that

$$B_r(x^J) = \sum_{T\in\mathcal{P}_{=r}([n]):J\subseteq T} \sum_{I:J\subseteq I\subseteq T} x^I(1-x)^{T\setminus I} = \sum_{T\in\mathcal{P}_{=r}([n]):J\subseteq T} x^J = \binom{n-|J|}{r-|J|} x^J$$

for any $J \subseteq [n]$. Hence, the Bernstein approximation of $f = \sum_{J\subseteq[n]} f_J x^J$ reads

$$B_r(f) = \sum_{J:J\subseteq[n],|J|\leq r} f_J \binom{n-|J|}{r-|J|} x^J. \tag{6.5}$$

Now we divide throughout by $\binom{n-1}{r-1}$ and add to both sides of (6.5) the quantity $\sum_J f_J\lambda_J x^J$ to get

$$\frac{B_r(f)}{\binom{n-1}{r-1}} + \sum_J f_J\lambda_J x^J = f.$$

As $B_r(1) = \binom{n}{r} = \frac{n}{r}\binom{n-1}{r-1}$, this gives $\frac{n}{r} f_{\max,\mathbf{Q}_n} = \frac{B_r(f_{\max,\mathbf{Q}_n})}{\binom{n-1}{r-1}}$ and thus we obtain

$$\frac{n}{r} f_{\max,\mathbf{Q}_n} - f = \frac{B_r(f_{\max,\mathbf{Q}_n} - f)}{\binom{n-1}{r-1}} - \sum_J \lambda_J f_J x^J. \tag{6.6}$$

As the polynomial $f_{\max,\mathbf{Q}_n} - f$ is nonnegative over $\{0,1\}^n$, it follows from the definition of the Bernstein operator that

$$B_r(f_{\max,\mathbf{Q}_n} - f) = \sum_{T\in\mathcal{P}_{=r}([n])} \sum_{I\subseteq T} (f_{\max,\mathbf{Q}_n} - f(\chi^I)) x^I (1-x)^{T\setminus I} \in H_r.$$

As $\lambda_J \geq 0$ for all $J$, after moving the terms $f_J\lambda_J x^J$ with $f_J > 0$ to the left hand side of (6.6), we obtain the claimed inequalities. □

### 6.1.3 The maximum stable set problem

Let $G = (V, E)$ be a graph and let $w \in \mathbb{R}_+^{|V|}$ be weights assigned to the nodes of $G$. The *maximum stable set problem* is to determine the maximum weight $w(S) = \sum_{i\in S} w_i$ of a stable set $S$ in $G$, called the weighted *stability number* of $(G, w)$ and denoted as $\alpha(G, w)$. Let $\mathrm{ST}(G)$ denote the polytope in $\mathbb{R}^{|V|}$, defined as the convex hull of the characteristic vectors of the stable sets of $G$:

$$\mathrm{ST}(G) := \mathrm{conv}\{\chi^S : S \subseteq V, \quad S \text{ is a stable set in } G\},$$

called the *stable set polytope* of G. Hence, computing $\alpha(G,w)$ is a linear optimization problem over the stable set polytope:

$$\alpha(G,w) = \max_{x \in \mathrm{ST}(G)} \sum_{i \in V} w_i x_i.$$

It is well known that computing $\alpha(G,w)$ is an NP-hard problem, already in the unweighted case when $w = \mathbf{e}$ [41]. An obvious linear relaxation of $\mathrm{ST}(G)$ is the *fractional stable set polytope* $\mathrm{FR}(G)$, defined as

$$\mathrm{FR}(G) := \{x \in \mathbb{R}^{|V|} : x \geq 0,\ x_i + x_j \leq 1\ \forall ij \in E\}.$$

By maximizing the linear objective function $w^T x$ over $\mathrm{FR}(G)$ we obtain an upper bound for the stability number:

$$\alpha^*(G,w) := \max_{x \in \mathrm{FR}(G)} \sum_{i \in V} w_i x_i, \tag{6.7}$$

called the *fractional stability number*.

We now consider another formulation for $\alpha(G,w)$ obtained by maximizing a suitable quadratic polynomial over the hypercube. Given node weights $w \in \mathbb{R}^{|V|}_+$, we consider edge weights $w_{ij}$ for the edges of $G$ satisfying the condition

$$w_{ij} \geq \min\{w_i, w_j\} \quad \text{for all edges } ij \in E. \tag{6.8}$$

For some of our results we will need to make a stronger assumption on the edge weights:

$$w_{ij} \geq \max\{w_i, w_j\} \quad \text{for all edges } ij \in E. \tag{6.9}$$

In the weighted case, unless specified otherwise, we will assume that the edge weights satisfy the weakest condition (6.8). In the unweighted case (i.e. $w_i = 1$ for all nodes $i \in V$), we simply define $w_{ij} = 1$ for all edges $ij \in E$. Once the edge weights are specified we define the (square-free quadratic) polynomials

$$p_{G,w} := \sum_{i \in V} w_i x_i - \sum_{ij \in E} w_{ij} x_i x_j,$$

$$f_{G,w} := \alpha(G,w) - p_{G,w} = \alpha(G,w) - \sum_{i \in V} w_i x_i + \sum_{ij \in E} w_{ij} x_i x_j. \tag{6.10}$$

In the unweighted case $p_{G,w}$ is the polynomial used earlier in the formulation (6.1).

In this chapter we are interested in establishing positivity certificates for the polynomial $f_{G,w}$ and in understanding what is the smallest integer $r$ for which $f_{G,w}$ belongs to the Handelman set $\mathcal{H}_r$, see Definition 6.10 below. It is clear that we get stronger positivity certificates if we can show that $f_{G,w} \in \mathcal{H}_r$ for lower values of the edge weights. This motivates our distinction between the above two conditions (6.8) and (6.9) on the edge weights.

Park and Hong [78] give the following reformulation for the maximum stable set problem (choosing $w_{ij} = \max\{w_i, w_j\}$ for the edge weights), we give a proof for completeness.

**Proposition 6.9.** *Given node weights $w \in \mathbb{R}_+^{|V|}$ and edge weights satisfying (6.8), the maximum stable set problem can be reformulated as*

$$\alpha(G,w) = \max_{x \in [0,1]^V} p_{G,w}(x) = \max_{x \in \{0,1\}^n} p_{G,w}(x). \tag{6.11}$$

**Proof.** As $p_{G,w}$ is square-free, it takes the same maximum value on $[0,1]^n$ and $\{0,1\}^n$. Clearly, the maximum value over $\{0,1\}^n$ is at least $\alpha(G,w)$ since $p_{G,w}$ evaluated at the characteristic vector of a maximum weight stable set is equal to $\alpha(G,w)$. It suffices now to observe that the maximum value of $p_{G,w}$ over $\{0,1\}^n$ is attained at the characteristic vector of a stable set. Indeed, for $S \subseteq V$, $p_{G,w}(\chi^S) = \sum_{i \in S} w_i - \sum_{ij \in E: i,j \in S} w_{ij}$. If $ij$ is an edge contained in $S$ with $w_j \geq w_i$, then $p_{G,w}(\chi^{S\setminus\{i\}}) - p_{G,w}(\chi^S) \geq w_{ij} - w_i \geq 0$. Hence we can replace $S$ by $S\setminus\{i\}$ without decreasing the objective value $p_{G,w}$. Iterating, we obtain that the maximum value of $p$ over $\{0,1\}^n$ is attained at a stable set. $\square$

By Proposition 6.1, the Handelman bound of order $r$ for problem (6.11) reads:

$$\bar{p}_{\text{han}}^{(r)}(G,w) := \inf\{\lambda : \lambda - p_{G,w} \in H_r\} \tag{6.12}$$

and, by Theorem 6.7, it satisfies the inequality: $\bar{p}_{\text{han}}^{(r)}(G,w) \leq \frac{n}{r}\alpha(G,w)$.

**Definition 6.10.** *We let* $\text{rk}_\text{H}(G,w)$ *denote the smallest integer $r$ for which*

$$\bar{p}_{\text{han}}^{(r)}(G,w) = \alpha(G,w),$$

*called the* Handelman rank *of the weighted graph* $(G,w)$*. Equivalently,* $\text{rk}_\text{H}(G,w)$ *is the smallest integer $r$ for which $f_{G,w}$ (in* (6.10)*) belongs to the Handelan set $\mathcal{H}_r$.*

For the all-ones weight function $w = \mathbf{e}$ (i.e., the unweighted case) we omit the subscript $w$ and simply write $p_G$, $f_G$, $\bar{p}_{\text{han}}^{(r)}(G)$, and $\text{rk}_\text{H}(G)$.

If $G$ has no edge then $\mathrm{rk}_{\mathrm{H}}(G, w) = 1$, since $\alpha(G, w) - p_{G,w} = \sum_{i \in V} w_i(1 - x_i) \in H_1$, and the Handelman rank is at least 2 if $G$ has at least one edge. As another example, it follows from Corollary 6.5 that, for the complete graph $K_n$, the polynomial $f_{K_n}$ belongs to $H_n$.

**Lemma 6.11.** *[78] The polynomial $f_{K_n} = \alpha(K_n) - p_{K_n} = 1 - \sum_{i=1}^{n} x_i + \sum_{1 \le i < j \le n} x_i x_j$ belongs to $H_n$.*

# 6.2 Handelman rank

## 6.2.1 Links to clique covers

In this section we show an upper bound for the Handelman bound in terms of fractional clique covers, and we characterize the graphs with Handelman rank at most 2.

First, we introduce fractional clique covers. Let $(G, w)$ be a weighted graph. A *fractional clique cover* of $(G, w)$ is a collection of cliques $C$ of $G$ together with scalars $\lambda_C \ge 0$ satisfying $\sum_C \lambda_C \chi^C = w$. Then the minimum value of $\sum_C \lambda_C$ is known as the weighted *fractional chromatic number* of $\overline{G}$:

$$\chi^*(\overline{G}, w) = \min \left\{ \sum_C \lambda_C : \sum_C \lambda_C \chi^C = w, \ \lambda_C \ge 0 \ \forall C \text{ clique of } G \right\}. \tag{6.13}$$

Note that if in addition we require the $\lambda_C$'s to be integer valued in (6.13) then we obtain the chromatic number $\chi(\overline{G}, w)$. Restricting to covers by cliques of size at most some given integer $r \ge 1$, we can define the parameter

$$\rho_r(G, w) := \min \left\{ \sum_C \lambda_C : \sum_C \lambda_C \chi^C = w, \ \lambda_C \ge 0 \ \forall C \text{ clique of } G \text{ with } |C| \le r \right\}, \tag{6.14}$$

which we call the *fractional $r$-clique cover number* of $(G, w)$. Thus

$$\rho_r(G, w) = \chi^*(\overline{G}, w) \ \text{ if } \ r \ge \omega(G),$$

where $\omega(G)$ denotes the largest size of a clique in $G$. In addition,

$$\rho_r(G, w) \ge \chi^*(\overline{G}, w) \ge \alpha(G, w).$$

As is well known, in relation (6.13) one can relax without loss of generality the equality $\sum_C \lambda_C \chi^C = w$ to the inequality $\sum_C \lambda_C \chi^C \ge w$. This extends to the fractional clique cover number. We include a short argument for clarity.

**Lemma 6.12.** *The parameter $\rho_r(G,w)$ from (6.14) is equal to the optimal value of the following program:*

$$\min\left\{\sum_C \lambda_C : \sum_C \lambda_C \chi^C \geq w,\ \lambda_C \geq 0\ \forall C \text{ clique of } G \text{ with } |C| \leq r\right\}. \tag{6.15}$$

**Proof.** Comparing (6.14) and (6.15), one only needs to show that the optimal value of (6.15) is at least $\rho_r(G,w)$. The argument is easier by looking at the dual linear programs. The dual of (6.14) reads

$$\max\left\{\sum_{i\in V} w_i x_i : \sum_{i\in C} x_i \leq 1\ \forall C \text{ clique of } G \text{ with } |C| \leq r\right\} \tag{6.16}$$

and the dual of (6.15) reads

$$\max\left\{\sum_{i\in V} w_i x_i : \sum_{i\in C} x_i \leq 1\ \forall C \text{ clique of } G \text{ with } |C| \leq r,\ x_i \geq 0\ \forall i \in V\right\}. \tag{6.17}$$

Suppose $x^* \in \mathbb{R}^n$ is an optimal solution of the program (6.16). Then define $y \in \mathbb{R}^n$ by setting $y_i = x_i$ if $x_i \geq 0$ and $y_i = 0$ otherwise. Then, $\sum_i w_i x_i^* \leq \sum_i w_i y_i$. It suffices now to show that $y$ is feasible for the program (6.17). For this, pick a clique $C$ with $|C| \leq r$, and let $C^*$ denote the subset of $C$ consisting of all elements $i \in C$ with $x_i^* \geq 0$. Then $C^*$ is again a clique with $|C^*| \leq r$ and thus $\sum_{i\in C^*} y_i = \sum_{i\in C^*} x_i^* \leq 1$, which concludes the proof. $\square$

For $r = 2$, $\rho_2(G,w)$ is the fractional edge cover number, which coincides with the fractional stability number $\alpha^*(G,w)$ of (6.7). Indeed, for $r = 2$, the program (6.7) coincides with (6.17) which is the dual of the program (6.15) defining $\rho_2(G,w)$.

**Proposition 6.13.** *Consider a weighted graph $(G,w)$ with edge weights satisfying (6.8). For any integer $r \geq 2$,*

$$\rho_r(G,w) - p_{G,w} \in H_r \quad \textit{and} \quad \bar{p}^{(r)}_{\text{han}}(G,w) \leq \rho_r(G,w).$$

**Proof.** Set $k = \rho_r(G,w)$. By definition (6.14), there exist scalars $\lambda_C \geq 0$ indexed by cliques $C$ of size at most $r$ such that (a) $\sum_C \lambda_C = k$, and (b) $w = \sum_C \lambda_C \chi^C$, i.e., $w_i = \sum_{C:i\in C} \lambda_C$ for all $i \in V$. In particular, this implies that (c) $\sum_{C:i,j\in C} \lambda_C \leq \min\{w_i, w_j\} \leq w_{ij}$ for all $ij \in E$. Moreover, by taking the inner product of both sides of (b) with the vector $(x_1, \cdots, x_n)^T$, we get $\sum_{i=1}^n w_i x_i = \sum_C \lambda_C x(C)$. Therefore,

$$\begin{aligned}
k - p_{G,w} &= \sum_C \lambda_C - \sum_{i\in V} w_i x_i + \sum_{ij\in E} w_{ij} x_i x_j \\
&= \sum_C \lambda_C \left(1 - \sum_{i\in C} x_i + \sum_{i<j:i,j\in C} x_i x_j\right) + \sum_{ij\in E} w_{ij} x_i x_j - \sum_C \lambda_C \sum_{i<j:i,j\in C} x_i x_j \\
&= \sum_C \lambda_C f_C + \sum_{ij\in E} w_{ij} x_i x_j - \sum_C \lambda_C \sum_{i<j:i,j\in C} x_i x_j,
\end{aligned}$$

setting $f_C = 1 - \sum_{i\in C} x_i + \sum_{i<j:i,j\in C} x_i x_j$. By Lemma 6.11, each $f_C$ lies in $H_r$ and thus the first sum lies in $H_r$. We now consider the remaining part:

$$\sum_{ij\in E} w_{ij} x_i x_j - \sum_C \lambda_C \sum_{i<j:i,j\in C} x_i x_j = \sum_{ij\in E} x_i x_j \left(w_{ij} - \sum_{C:i,j\in C} \lambda_C\right),$$

which belongs to $H_2$ since the scalars $w_{ij} - \sum_{C:i,j\in C} \lambda_C$ are nonnegative by (c). Thus we have shown that $k - p_{G,w} \in H_r$, which gives directly $\bar{p}^{(r)}_{\text{han}}(G, w) \le k$. □

Next, we show that equality $\bar{p}^{(r)}_{\text{han}}(G, w) = \rho_r(G, w)$ holds for $r = 2$. Note that for $r \ge 3$, the strict inequality $\bar{p}^{(r)}_{\text{han}}(G, w) < \rho_r(G, w)$ is possible. For instance, for the odd circuit $C_{2n+1}$, $\bar{p}^{(3)}_{\text{han}}(C_{2n+1}) = \alpha(C_{2n+1}) < \rho_3(C_{2n+1}) = \alpha^*(C_{2n+1})$ holds (see Proposition 6.28 below).

**Theorem 6.14.** *Consider a weighted graph $(G, w)$ with edge weights satisfying (6.8). Then, $\bar{p}^{(2)}_{\text{han}}(G, w) = \rho_2(G, w)$.*

**Proof.** Set $k = \bar{p}^{(2)}_{\text{han}}(G, w)$. In what follows we construct a fractional 2-clique covering of $(G, w)$ of value $k$, which shows the inequality $\rho_2(G, w) \le \bar{p}^{(2)}_{\text{han}}(G, w)$ and concludes the proof. By assumption, the polynomial $k - p_{G,w}$ belongs to $H_2$ and thus has a decomposition:

$$k - p_{G,w} = \sum_{ij\in E_n} a_{ij}(1-x_i)(1-x_j) + b_{ij} x_i (1-x_j) + c_{ij} x_j (1-x_i) + d_{ij} x_i x_j \quad (6.18)$$

where all scalars $a_{ij}, b_{ij}, c_{ij}, d_{ij} \ge 0$ and $E_n$ denotes the set of ordered pairs $ij$ with $1 \le i < j \le n$. By evaluating the coefficients of the monomials 1, $x_i$ and $x_i x_j$ we get the relations:

$$k = \sum_{ij\in E_n} a_{ij},$$

$$-w_i = -\sum_{j:j>i} a_{ij} - \sum_{j:j<i} a_{ji} + \sum_{j:j>i} b_{ij} + \sum_{j:j<i} c_{ji} \quad \text{for any } i \in V,$$

$$a_{ij} - b_{ij} - c_{ij} + d_{ij} = \begin{cases} w_{ij} & \text{if } ij \in E \\ 0 & \text{otherwise.} \end{cases} \quad \text{for any pair } ij \in E_n. \tag{6.19}$$

First we observe that we can find another decomposition of $k - p_{G,w}$, of the form (6.20) below, which involves quadratic terms only for the edges of $G$ but has additional linear terms. For any pair $ij \in E_n$, set

$$f_{ij} = a_{ij}(1-x_i)(1-x_j) + b_{ij}x_i(1-x_j) + c_{ij}x_j(1-x_i) + d_{ij}x_ix_j$$

so that the decomposition (6.18) reads: $k - p_{G,w} = \sum_{ij \in E_n} f_{ij}$. We now show that, for any $ij \in E_n \setminus E$, the polynomial $f_{ij}$ belongs to $H_1$. Indeed, pick a pair $ij$ which is not an edge. By (6.19), we have: $d_{ij} = b_{ij} + c_{ij} - a_{ij}$, so that we can rewrite $f_{ij}$ as

$$f_{ij} = x_i(b_{ij} - a_{ij}) + x_j(c_{ij} - a_{ij}) + a_{ij}.$$

We distinguish several cases:

- If $b_{ij} - a_{ij} \geq 0$ and $c_{ij} - a_{ij} \geq 0$ then we get a representation in $H_1$ for $f_{ij}$.
- If $b_{ij} - a_{ij} \leq 0$ and $c_{ij} - a_{ij} \geq 0$ then rewrite $f_{ij}$ as:

$$f_{ij} = (1-x_i)(a_{ij} - b_{ij}) + x_j(c_{ij} - a_{ij}) + b_{ij} \in H_1.$$

- Analogously if $b_{ij} - a_{ij} \geq 0$ and $c_{ij} - a_{ij} \leq 0$.
- If $b_{ij} - a_{ij} \leq 0$ and $c_{ij} - a_{ij} \leq 0$ then rewrite $f_{ij}$ as:

$$f_{ij} = (1-x_i)(a_{ij} - b_{ij}) + (1-x_j)(a_{ij} - c_{ij}) + b_{ij} + c_{ij} - a_{ij}$$

which is again a representation in $H_1$ since $b_{ij} + c_{ij} - a_{ij} = d_{ij} \geq 0$. Hence, we have shown $f_{ij} \in H_1$ for all nonedges and thus we obtain a new representation of $k - p_{G,w}$ of the form:

$$\begin{aligned} k - p_{G,w} &= \sum_{ij \in E} a_{ij}(1-x_i)(1-x_j) + b_{ij}x_i(1-x_j) + c_{ij}x_j(1-x_i) + d_{ij}x_ix_j \\ &+ \sum_{i \in V} f_ix_i + g_i(1-x_i), \end{aligned} \tag{6.20}$$

where all coefficients $a_{ij}, b_{ij}, c_{ij}, d_{ij}, f_i, g_i$ are nonnegative scalars. Then, we obtain:

$$k = \sum_{ij \in E} a_{ij} + \sum_{i \in V} g_i, \tag{6.21}$$

and for all $i \in V$:

$$-w_i = -\sum_{j:j>i,ij\in E} a_{ij} - \sum_{j:j<i,ij\in E} a_{ji} + \sum_{j:j>i,ij\in E} b_{ij} + \sum_{j:j<i,ij\in E} c_{ji} + f_i - g_i. \quad (6.22)$$

We now build a fractional clique cover. For this consider the vector:

$$u = \sum_{ij\in E, i<j} a_{ij}\chi^{\{i,j\}} + \sum_{i\in V} g_i\chi^{\{i\}}.$$

We check that $u_i \geq w_i$ for all $i \in V$. For this fix $i$ and set $N = \{j : ij \in E\}$. We have:

$$u_i = \sum_{j\in N:j>i} a_{ij} + \sum_{j\in N:j<i} a_{ji} + g_i.$$

Using (6.22) we get:

$$w_i = \sum_{j\in N:j>i} a_{ij} + \sum_{j\in N:j<i} a_{ji} - \sum_{j\in N:j>i} b_{ij} - \sum_{j\in N:j<i} c_{ji} - f_i + g_i.$$

Thus $u_i \geq w_i$ is equivalent to

$$0 \geq -\sum_{j\in N:j>i} b_{ij} - \sum_{j\in N:j<i} c_{ji} - f_i.$$

It suffices now to observe that indeed $f_i \geq 0$, $\sum_{j\in N:j>i} b_{ij} \geq 0$, and $\sum_{j\in N:j<i} c_{ji} \geq 0$. Hence $u$ is a fractional 2-clique cover of $(G, w)$ with value $\sum_{ij\in E} a_{ij} + \sum_{i\in V} g_i = k$ by (6.21). This implies that $\rho_2(G, w) \leq k$ and concludes the proof. □

Now we can characterize the graphs with Handelman rank equal to 2.

**Corollary 6.15.** *The Handelman bound of order 2 is exact if and only if there is a fractional edge covering of value $\alpha(G, w)$, i.e.,*

$$\bar{p}^{(2)}_{\text{han}}(G, w) = \alpha(G, w) \Longleftrightarrow \alpha(G, w) = \rho_2(G, w) \Longleftrightarrow \alpha^*(G, w) = \alpha(G, w).$$

It is well known that the equality $\alpha(G, w) = \alpha^*(G, w)$ holds for any node weights $w \in \mathbb{R}^{|V|}_+$ if and only if $G$ is bipartite [64, Section 4]. This implies that the Handelman rank of any weighted bipartite graph is at most 2, settling an open question of Park and Hong [78] who proved the result in the unweighted case.

**Corollary 6.16.** *If $G$ is bipartite, then $\mathrm{rk}_{\mathrm{H}}(G, w) \leq 2$ for any node weights $w \in \mathbb{R}^{|V|}_+$.*

On the other hand, the Handelman hierarchy is sometimes exact at order 2 for non-bipartite graphs, as the next example shows.

**Example 6.17.** *Let $G$ be the graph on $2t$ nodes obtained by taking the clique sum of $t$ copies of $K_{t+1}$ along a common clique $K_t$. Then $\alpha(G) = t$, $\rho_2(G) = t$ (since one can cover all nodes by $t$ disjoint edges), and thus the Handelman relaxation of order 2 is exact:* $\mathrm{rk}_{\mathrm{H}}(G) = 2$.

### 6.2.2 Bounds for the Handelman rank

In this section, we show some lower and upper bounds for the Handelman rank of weighted graphs. The upper bounds hold when assuming that the edge weights satisfy (6.9).

#### Lower bound

We start with the following lemma from [78, Prop. 3.3] which we prove for completeness.

**Lemma 6.18.** *Consider a square-free polynomial*

$$f(x) = a_0 + \sum_{i\in[n]} a_i x_i + \sum_{I\subseteq[n]:|I|\geq 2} a_I x^I.$$

*If $\lambda - f \in H_r$, then $\lambda - a_0 \geq \sum_{i\in[n]} a_i/r$.*

**Proof.** Say, $\lambda - f = \sum_{T\in\mathcal{P}_{=r}(V), I\subseteq T} c_{I,T} x^I (1-x)^{T\setminus I}$ with $c_{I,T} \geq 0$. Evaluating the constant term we find that

$$\lambda - a_0 = \sum_{T\in\mathcal{P}_{=r}(V)} c_{\emptyset,T}.$$

Evaluating the coefficient of $x_i$ we get:

$$-a_i = \sum_{T\in\mathcal{P}_{=r}(V):i\in T} \left(c_{\{i\},T} - c_{\emptyset,T}\right).$$

Summing up over all $i \in V = [n]$ gives:

$$\begin{aligned} -\sum_{i\in[n]} a_i &= \sum_{i\in[n]} \sum_{T\in\mathcal{P}_{=r}(V):i\in T} c_{\{i\},T} - \sum_{i\in[n]} \sum_{T\in\mathcal{P}_{=r}(V):i\in T} c_{\emptyset,T} \\ &\geq -\sum_{T\in\mathcal{P}_{=r}(V)} r c_{\emptyset,T} = -r(\lambda - a_0), \end{aligned}$$

which implies $\lambda - a_0 \geq \sum_{i\in[n]} a_i/r$. $\square$

Applying Lemma 6.18 to the polynomial $p_{G,w}$ we obtain the following lower bound on the Handelman rank.

**Proposition 6.19.** *Consider a weighted graph $(G, w)$ where the edge weights satisfy (6.8). Then, $\bar{p}_{\mathrm{han}}^{(r)}(G, w) \geq \frac{\sum_{i=1}^n w_i}{r}$. Therefore,*

$$\mathrm{rk}_{\mathrm{H}}(G, w) \geq \frac{\sum_{i=1}^n w_i}{\alpha(G, w)}. \tag{6.23}$$

For the unweighted complete graph $G = K_n$, the lower bound is equal to $n$, which implies $\mathrm{rk}_{\mathrm{H}}(K_n) \geq n$. Hence equality holds: $\mathrm{rk}_{\mathrm{H}}(K_n) = n$ and the lower bound is tight.

### The first upper bound

First we show an upper bound for the Handelman rank of a weighted graph $(G, w)$, in terms of parameters of the unweighted graph $G$.

**Theorem 6.20.** *Consider a weighted graph $(G, w)$ where the edge weights satisfy (6.9). Then,*

$$\mathrm{rk}_{\mathrm{H}}(G, w) \leq |V(G)| - \alpha(G) + 1. \tag{6.24}$$

Note that the upper bund (6.24) is tight for the unweighted complete graph $K_n$. The proof of Theorem 6.20 relies on Lemma 6.21 below which will allow to use induction on the number of nodes.

In what follows we use the following notation: Given a weighted graph $(G, w)$ and a subset $U \subseteq V$, $(G \backslash U, w)$ denotes the weighted graph $G \backslash U$ where the node and edge weights are obtained from those of $G$ simply by restricting to nodes and edges of $G \backslash U$. For a node $i \in V$, recall that we denote $G - i$ as the graph obtained by deleting node $i$ from $G$, and that we denote $G \ominus i$ as the graph obtained from $G$ by removing $i$ as well as its neighbours.

**Lemma 6.21.** *Consider a weighted graph $(G, w)$ where the edge weights satisfy (6.9). For any node $i \in V$, one has*

$$\mathrm{rk}_{\mathrm{H}}(G, w) \leq \max\{\mathrm{rk}_{\mathrm{H}}(G - i, w) + 1, \mathrm{rk}_{\mathrm{H}}(G \ominus i, w) + 1, 3\}.$$

**Proof.** Recall the polynomial $f_{G,w} = \alpha(G, w) - p_{G,w}$ from (6.10). For convenience we consider the node $i = n$ and we set $\underline{x} = (x_1, x_2, \ldots, x_{n-1})$ so that $x = (\underline{x}, x_n)$. By Lemma 6.6,

$$f_{G,w}(x) = (1 - x_n) f_{G,w}(\underline{x}, 0) + x_n f_{G,w}(\underline{x}, 1). \tag{6.25}$$

First, we can write $f_{G,w}(\underline{x},0) = f_{G-n,w}(\underline{x}) + g_1$, where $g_1 = \alpha(G,w) - \alpha(G-n,w) \geq 0$. Moreover, we have the identity $f_{G,w}(\underline{x},1) = f_{G\ominus n,w}(\underline{x}) + g_2(\underline{x})$, after setting

$$\begin{aligned} g_2(\underline{x}) &= \underbrace{\alpha(G,w) - w_n - \alpha(G \ominus n, w)}_{\geq 0} + \sum_{i \in N(n)} \underbrace{(w_{in} - w_i)}_{\geq 0} x_i \\ &+ \sum_{ij \in E(G-n) \setminus E(G \ominus n)} \underbrace{w_{ij}}_{\geq 0} x_i x_j \in H_2. \end{aligned}$$

Here we have used the assumption (6.9) in order to claim that $w_{in} \geq w_i$ for all $i \in N(n)$. Combining with (6.25), we obtain

$$f_{G,w}(x) = (1 - x_n) f_{G-n,w}(\underline{x}) + x_n f_{G\ominus n,w}(\underline{x}) + h(x),$$

where $h(x) = (1 - x_n) g_1 + x_n g_2(\underline{x}) \in H_3$. Hence the lemma is proved. $\square$

**Proof.** *(of Theorem 6.20)* We show (6.24) by induction on the number of nodes $|V(G)|$. If $G$ has no edge then $\mathrm{rk_H}(G,w) = 1$ and thus the result holds for $|V(G)| = 1$. If $\alpha(G) = |V| - 1$ then $G$ is bipartite and thus $\mathrm{rk_H}(G,w) = 2$ (by Corollary 6.16) and thus the result holds. Assume now that $|V(G)| \geq 2$ and $\alpha(G) \leq |V(G)| - 2$. Then there exists a node $i \in V$ satisfying

$$\alpha(G - i) = \alpha(G).$$

In particular, $i$ is adjacent to at least one node: $|N(i)| \geq 1$. Using the induction assumption for the graphs $G - i$ and $G \ominus i$, we obtain that

$$\begin{aligned} \mathrm{rk_H}(G - i, w) &\leq (|V(G)| - 1) - \alpha(G - i) + 1 = |V(G)| - \alpha(G - i) \\ &= |V(G)| - \alpha(G), \\ \mathrm{rk_H}(G \ominus i, w) &\leq (|V(G)| - |N(i)| - 1) - \alpha(G \ominus i) + 1 \\ &= |V(G)| - |N(i)| - \alpha(G \ominus i) \leq |V(G)| - \alpha(G). \end{aligned}$$

Here we have used the (easy to check) inequality $\alpha(G) \leq \alpha(G \ominus i) + |N(i)|$. Now we can use Lemma 6.21 and conclude that $\mathrm{rk_H}(G,w) \leq |V(G)| - \alpha(G) + 1$. $\square$

### The second upper bound

We now give another upper bound for the Handelman rank of a weighted graph $(G,w)$, which depends on the specific node weights. Consider an inequality $w^T x \leq b$

which is valid for $\mathrm{ST}(G)$, where we assume $w \in \mathbb{N}^V$ and $b \in \mathbb{N}$; obviously $b \geq \alpha(G, w)$. Define the *defect* of this inequality as

$$\mathrm{defect}_G(w, b) = 2(\alpha^*(G, w) - \min\{b, \alpha^*(G, w)\}). \tag{6.26}$$

Note that the defect is a nonnegative integer number, since the node weights $w$ are integer valued and there is a $\{0, 1/2, 1\}$-valued vector $x \in \mathrm{FR}(G)$ maximizing $w^T x$ over $\mathrm{FR}(G)$ (see [71, Section 2.c]). We have the following result on the polynomial $b - p_{G,w}$.

**Theorem 6.22.** *Assume $w^T x \leq b$ is valid for $\mathrm{ST}(G)$, where $w \in \mathbb{N}^V$ and $b \in \mathbb{N}$, and let the edge weights satisfy (6.9). Then the polynomial $b - p_{G,w}$ belongs to $H_{r+2}$, where $r = \mathrm{defect}_G(w, b)$ is defined in (6.26).*

The proof uses the result of Lovász and Schrijver [65] from Lemma 6.23 below. It is along the similar lines as their proof of [65, Theorem 2.13] where they upper bound the $N$-index of the inequality $w^T x \leq \alpha(G, w)$ by the quantity $2(\alpha^*(G, w) - \alpha(G, w))$. We return to the construction of Lovász and Schrijver [65] in Section 6.3.2.

**Lemma 6.23.** *[65, Lemma 2.12] Consider node weights $w \in \mathbb{N}^V$ for which*

$$\alpha(G, w) < \alpha^*(G, w).$$

*Then, there exists a node $i \in V$ such that every vector $x \in \mathrm{FR}(G)$ maximizing $w^T x$ over $\mathrm{FR}(G)$ (i.e., $w^T x = \alpha^*(G, w)$) satisfies $x_i = \frac{1}{2}$.*

**Proof.** *(of Theorem 6.22)* The proof is by induction on the defect $r := 2(\alpha^*(G, w) - \min\{b, \alpha^*(G, w)\})$. If $r = 0$, i.e., $b \geq \alpha^*(G, w) = \rho_2(G, w)$, then the result follows from Proposition 6.13, since $b - p_{G,w} = (b - \rho_2(G, w)) + (\rho_2(G, w) - p_{G,w}) \in H_2$.

Assume now that $b < \alpha^*(G, w)$ (i.e., $r > 0$). Then $\alpha(G, w) \leq b < \alpha^*(G, w)$ and thus Lemma 6.23 can be applied. Hence there exists one node, denoted as $n$ for convenience, such that every vector $x \in \mathrm{FR}(G)$ optimizing $w^T x$ over $\mathrm{FR}(G)$ has $x_n = 1/2$. This trivially implies $w_n > 0$. Let $w_{G-n}$ denote the restriction of $w$ to the nodeset of $G - n$ and define $w' \in \mathbb{R}^{|V|}$ which coincides with $w$ except $w'_n = 0$. Analogously, $w_{G \ominus n}$ denotes the restriction of $w$ to the nodeset of $G \ominus n$ and $w'' \in \mathbb{R}^{|V|}$ coincides with $w$ except $w''_i = 0$ if $i$ is equal or adjacent to $n$. Observe that $\alpha^*(G, w') = \alpha^*(G - n, w_{G-n})$ and $\alpha^*(G, w'') = \alpha^*(G \ominus n, w_{G \ominus n})$.

We consider the two inequalities $w_{G-n}^T x \leq b$ and $w_{G \ominus n}^T x \leq b - w_n$, which are clearly valid for $\mathrm{ST}(G - n)$ and $\mathrm{ST}(G \ominus n)$, respectively. Their defects are respectively denoted as

$$\begin{aligned} r' &= 2(\alpha^*(G - n, w_{G-n}) - \min\{b, \alpha^*(G - n, w_{G-n})\}) \\ &= 2(\alpha^*(G, w') - \min\{b, \alpha^*(G, w')\}) \end{aligned}$$

and

$$\begin{aligned} r'' &= 2(\alpha^*(G \ominus n, w_{G-n}) - \min\{b - w_n, \alpha^*(G \ominus n, w_{G-n})\}) \\ &= 2(\alpha^*(G, w'') - \min\{b - w_n, \alpha^*(G, w'')\}). \end{aligned}$$

We show that both defects smaller than $r$, i.e., that $r', r'' < r$.

First, we show that $r' < r$. This is clear if $b \geq \alpha^*(G, w')$ as then $r' = 0 < r$. Now, we can suppose that $b < \alpha^*(G, w')$ and it suffices to show that $\alpha^*(G, w') < \alpha^*(G, w)$. For this, let $y$ be a vertex of $\mathrm{FR}(G)$ maximizing $(w')^T x$ over $\mathrm{FR}(G)$. Then,

$$w^T y = (w')^T y + w_n y_n = \alpha^*(G, w') + w_n y_n \leq \alpha^*(G, w).$$

If $y_n > 0$, then $\alpha^*(G, w') \leq \alpha^*(G, w) - w_n y_n < \alpha^*(G, w)$, since $w_n > 0$. If $y_n = 0$ then, by Lemma 6.23, $y$ does not maximize $w^T x$ over $\mathrm{ST}(G)$ and thus $w^T y < \alpha^*(G, w)$, giving again $\alpha^*(G, w') < \alpha^*(G, w)$. Thus $r' < r$ holds.

We now show that $r'' < r$. This is clear if $b - w_n \geq \alpha^*(G, w'')$ as then $r'' = 0 < r$. Now, we can suppose that $b - w_n < \alpha^*(G, w'')$ and it suffices to show that $\alpha^*(G, w'') + w_n < \alpha^*(G, w)$. For this let $z$ be a vertex of $\mathrm{FR}(G)$ maximizing $(w'')^T x$ over $\mathrm{FR}(G)$. Define the new vector $\bar{z} \in \mathbb{R}^{|V|}$ which coincides with $z$ except $\bar{z}_n = 1$ and $\bar{z}_i = 0$ if $i$ is adjacent to $n$. Then, $\bar{z} \in \mathrm{FR}(G)$ and $w^T \bar{z} = (w'')^T z + w_n = \alpha^*(G, w'') + w_n$. As $\bar{z}_n \neq \frac{1}{2}$, we deduce from Lemma 6.23 that $w^T \bar{z} < \alpha^*(G, w)$ thus showing $\alpha^*(G, w'') + w_n < \alpha^*(G, w)$.

Thus $r' + 2, r'' + 2 \leq r + 1$ and using the induction assumption we can conclude that the following two polynomials both lie in the Handelman set of order $r + 1$:

$$\begin{aligned} f_1 &= b - \sum_{i \in V(G-n)} w_i x_i + \sum_{ij \in E(G-n)} w_{ij} x_i x_j \in H_{r+1}, \\ f_2 &= b - w_n - \sum_{i \in V(G \ominus n)} w_i x_i + \sum_{ij \in E(G \ominus n)} w_{ij} x_i x_j \in H_{r+1}. \end{aligned}$$

Define $f := b - p_{G,w}$ and observe that

$$f(\underline{x}, 0) = f_1 \text{ and } f(\underline{x}, 1) = f_2 + \sum_{i \in N(n)} (w_{in} - w_i) x_i + \sum_{ij \in E(G-n) \setminus E(G \ominus n)} w_{ij} x_i x_j.$$

By Lemma 6.6, $f(x) = (1 - x_n) f(\underline{x}, 0) + x_n f(\underline{x}, 1)$, thus implying $f \in H_{r+2}$. □

Considering that the defect of $w^T x \leq \alpha(G, w)$ is $2(\alpha^*(G, w) - \alpha(G, w))$, by Theorem 6.22 we have the following upper bound for $\mathrm{rk}_{\mathrm{H}}(G, w)$.

**Corollary 6.24.** *Consider a weighted graph $(G, w)$ with integer node weights $w \in \mathbb{N}^V$ and where the edge weights satisfy (6.9). Then,*

$$\mathrm{rk_H}(G, w) \leq 2(\alpha^*(G, w) - \alpha(G, w)) + 2. \tag{6.27}$$

**Remark 6.25.** *The upper bound (6.24) holds for any weight function $w \in \mathbb{R}_+^{|V|}$, while the upper bound (6.27) holds for any integral weight function $w \in \mathbb{N}^V$ (which can be assumed without loss of generality). It turns out that these two upper bounds are not comparable. Indeed, for the unweighted odd circuit $C_{2n+1}$, (6.24) and (6.27) give $n + 2$ and 3, respectively. On the other hand, consider an unweighted graph consisting of $n$ isolated nodes, then (6.24) and (6.27) read 1 and 2, respectively.*

### 6.2.3 Handelman ranks of some special classes of graphs

As an application we can now determine the Handelman rank of some special classes of graphs, including perfect graphs, odd circuits and their complements.

#### Perfect graphs

A graph $G$ is said to be *perfect* if equality $\omega(H) = \chi(H)$ holds for all induced subgraphs $H$ of $G$ (including $H = G$). We will use the following properties of perfect graphs and refer to [63] for details. If $G$ is perfect then its complement $\overline{G}$ is perfect as well and thus $\alpha(H) = \chi(\overline{H})$ for all induced subgraphs $H$ of $G$. Moreover, $\alpha(G, w) = \chi(\overline{G}, w)$ for any node weights $w \in \mathbb{R}_+^{|V|}$. We also use the following well-known fact: For any graph $G$, $|V(G)| \leq \alpha(G)\chi(G)$, with equality if $G$ is perfect and vertex transitive (see, e.g. [88, Section 67.4]). We can show the following upper bound for the Handelman rank of weighted perfect graphs.

**Proposition 6.26.** *Consider a weighted graph $(G, w)$ where the edge weights satisfy (6.9). If $G$ is perfect then $\mathrm{rk_H}(G, w) \leq \omega(G)$. Moreover, in the unweighted case, $\mathrm{rk_H}(G) = \omega(G)$ if $G$ is vertex-transitive.*

**Proof.** We know from Proposition 6.13 that $\chi(\overline{G}, w) - p_{G,w} \in H_{\omega(G)}$. As $G$ is perfect, $\alpha(G, w) = \chi(\overline{G}, w)$ and thus $\alpha(G, w) - p_{G,w} \in H_{\omega(G)}$, which shows $\mathrm{rk_H}(G, w) \leq \omega(G)$. Assume now that $w$ is the all-ones vector and that $G$ is perfect and vertex-transitive. Then, we have equality: $|V(G)| = \alpha(G)\chi(G) = \alpha(G)\omega(G)$. Using Proposition 6.19, we obtain that $\mathrm{rk_H}(G) \geq |V(G)|/\alpha(G) = \omega(G)$, which implies $\mathrm{rk_H}(G) = \omega(G)$. □

**Remark 6.27.** *The inequality $\mathrm{rk_H}(G) \leq \omega(G)$ can be strict for some perfect graphs. This is the case, for instance, for the graph $G$ from Example 6.17, which is perfect with $\omega(G) = t + 1$ and $\mathrm{rk_H}(G) = 2$. Figure 6.1 shows this graph for the case $t = 2$.*

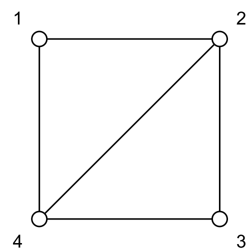


Figure 6.1: Example: a perfect graph

**Odd circuits and their complements**

Park and Hong [78] show that the Handelman rank of an odd circuit is equal to 3. Here we show that the Handelman rank of a weighted odd circuit is at most 3, answering an open question of [78], and we also consider the Handelman rank of complements of odd circuits.

**Proposition 6.28.** *Consider a weighted odd circuit* $(C_{2n+1}, w)$ *and its complement* $(\overline{C_{2n+1}}, w)$*, where the edge weights satisfy (6.9). Then,*

$$\mathrm{rk_H}(C_{2n+1}, w) \leq 3 \quad \textit{and} \quad \mathrm{rk_H}(\overline{C_{2n+1}}, w) \leq n+1.$$

*Moreover, equality holds in the unweighted case:* $\mathrm{rk_H}(C_{2n+1}) = 3$ *and* $\mathrm{rk_H}(\overline{C_{2n+1}}) = n+1$.

**Proof.** For any node $i$, both graphs $C_{2n+1} - i$ and $C_{2n+1} \ominus i$ are bipartite and thus $\mathrm{rk_H}(C_{2n+1} - i, w)$, $\mathrm{rk_H}(C_{2n+1} \ominus i, w) \leq 2$ by Corollary 6.16. Applying Lemma 6.21, we obtain that $\mathrm{rk_H}(C_{2n+1}, w) \leq 3$. Similarly, for any node $i$, both graphs $\overline{C_{2n+1}} - i$ and $\overline{C_{2n+1}} \ominus i$ are perfect with clique number at most $n$ and thus, from Proposition 6.26, $\mathrm{rk_H}(\overline{C_{2n+1}} - i, w)$, $\mathrm{rk_H}(\overline{C_{2n+1}} \ominus i, w) \leq n$. Applying again Lemma 6.21 we deduce that $\mathrm{rk_H}(\overline{C_{2n+1}}, w) \leq n+1$. In the unweighted case, the lower bounds $\mathrm{rk_H}(C_{2n+1}) \geq 3$ and $\mathrm{rk_H}(\overline{C_{2n+1}}) \geq n+1$ follow from Proposition 6.19. Indeed, $\mathrm{rk_H}(C_{2n+1}) \geq \frac{2n+1}{\alpha(C_{2n+1})} = \frac{2n+1}{n} > 2$ and $\mathrm{rk_H}(\overline{C_{2n+1}}) \geq \frac{2n+1}{\alpha(\overline{C_{2n+1}})} = \frac{2n+1}{2} > n$. $\square$

As an application we obtain the following characterization of perfect graphs, which is in the same spirit as the following well-known characterization due to Lovász [63]: $G$ is perfect if and only if $|V(H)| \leq \alpha(H)\omega(H)$ for all induced subgraphs $H$ of $G$.

**Corollary 6.29.** *A graph* $G$ *is perfect if and only if* $\mathrm{rk_H}(H) \leq \omega(H)$ *for every induced subgraph* $H$ *of* $G$.

**Proof.** The 'only if' part follows from Proposition 6.26. Conversely, assume that $G$ is not perfect. Using the 'strong perfect graph theorem' of Chudnovsky, Robertson, Seymour and Thomas [12], we know that $G$ contains an induced subgraph $H$ which is an odd circuit or its complement. By Proposition 6.28, $\mathrm{rk_H}(H) = \chi(H) > \omega(H)$, concluding the proof. $\square$

**Remark 6.30.** *As noted earlier, the upper bound 3 for the Handelman rank of an odd circuit also follows from the upper bound from Corollary 6.24 in terms of the defect. Indeed, $\alpha^*(C_{2n+1}) = (2n+1)/2$, so that the defect of the inequality*

$$\sum_{i \in V(C_{2n+1})} x_i \leq n = \alpha(C_{2n+1})$$

*is equal to $2((2n+1)/2 - n) = 1$ and thus relation (6.27) gives the upper bound 3.*

*Park and Hong [78] show that the Handelman rank of an odd circuit is at most 3 by constructing an explicit decomposition of the polynomial $\alpha(C_{2n+1}) - p_{C_{2n+1}}$ in the Handelman set $H_3$. We illustrate their argument for the case of $C_5$, see Figure 6.2. For the graph $C_5$, we have:*

$$\alpha(C_5) - p_{C_5} = 2 - \sum_{i=1}^{5} x_i + \sum_{i=1}^{4} x_i x_{i+1} + x_1 x_5 = f_{123} + f_{145} + f'_{1,34},$$

*where*

$$f_{123} = 1-(x_1+x_2+x_3)+x_1x_2+x_1x_3+x_2x_3 = (1-x_1)(1-x_2)(1-x_3)+x_1x_2x_3 \in H_3,$$

$$f_{145} = 1-(x_1+x_4+x_5)+x_1x_4+x_1x_5+x_4x_5 = (1-x_1)(1-x_4)(1-x_5)+x_1x_4x_5 \in H_3,$$

$$\begin{aligned} f'_{1,34} &= f_{134}(1-x_1, x_3, x_4) = x_1 - x_1x_3 - x_1x_4 + x_3x_4 \\ &= x_1(1-x_3)(1-x_4) + (1-x_1)x_3x_4 \in H_3. \end{aligned}$$

*In the above decomposition, $f_{123}$ and $f_{145}$ are the polynomials corresponding to the two cliques $\{1,2,3\}$ and $\{1,4,5\}$ (obtained by adding the edges 13 and 14 to $C_5$), and the polynomial $f'_{1,34}$ permits to cancel the quadratic terms $x_1x_3$ and $x_1x_4$ corresponding to the added edges 13 and 14 and to add the quadratic term $x_3x_4$. This construction extends easily to an arbitrary odd circuit, showing $\mathrm{rk_H}(C_{2n+1}) \leq 3$.*

We conclude with bounding the Handelman rank of two more classes of graphs.

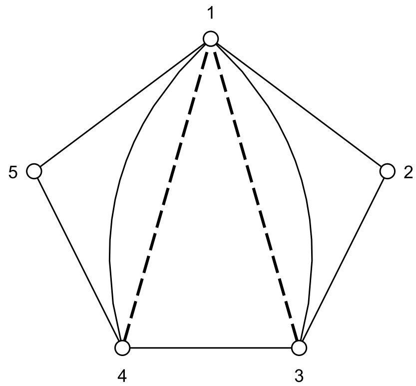


Figure 6.2: Odd circuit $C_5$

**Example 6.31.** *Consider the odd wheel $W_{2n+1}$, which is the graph obtained from an odd circuit $C_{2n+1}$ by adding a new node (the apex node, denoted as $v_0$) and making it adjacent to all nodes of $C_{2n+1}$. Since by deleting the apex node $v_0$ one obtains $C_{2n+1}$ with Handelman rank 3, Lemma 6.21 implies that the Handelman rank of the wheel $W_{2n+1}$ is at most 4; note that this bound also holds for any weighted wheel. Moreover, the complement of $W_{2n+1}$ has the same Handelman rank as the complement of $C_{2n+1}$ (since node $v_0$ is isolated, and apply Lemma 6.34 (iv) below).*

**Example 6.32.** *We now consider the graphs $G_k$, constructed by Lipták and Tuncel [62] and defined as in Figure 6.3. Hence, for $k = 2$, $G_2$ is the circuit $C_5$ with a new node adjacent to three consecutive nodes of $C_5$. We show that, for any $k \geq 2$, the Handelman rank of the graph $G_k$ is equal to 3 or 4.*

*As $G_k$ has $3k$ nodes and $\alpha(G_k) = k$, the lower bound (6.23) for the Handelman rank gives $\mathrm{rk}_{\mathrm{H}}(G_k) \geq 3$. Now, we look at the upper bound for the Handelman rank. First, we consider the case $k = 2$. As in Remark 6.30, we can give an explicit decomposition for the polynomial $\alpha(G_2) - p_{G_2}$. Namely,*

$$\alpha(G_2) - p_{G_2} = f_{1234} + f_{456} + f'_{4,36},$$

*where*

$$f_{1234} = 1 - \sum_{i=1}^{4} x_i + \sum_{1 \leq i < j \leq 4} x_i x_j \in H_4,$$

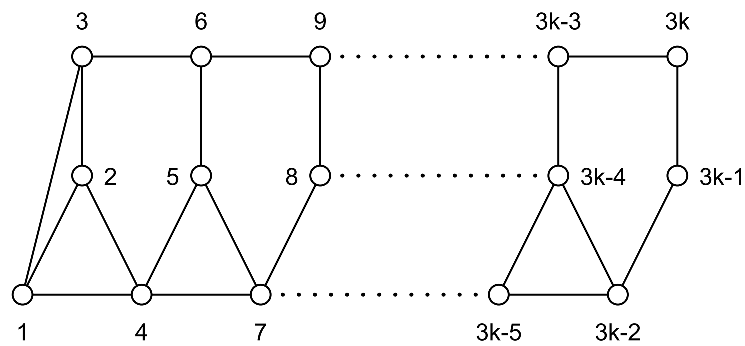


Figure 6.3: Graph $G_k$

$$f_{456} = 1 - \sum_{i=4}^{6} x_i + (x_4x_5 + x_4x_6 + x_5x_6) \in H_3,$$

$$f'_{4,36} = f_{436}(1 - x_4, x_3, x_6) = x_4(1 - x_3)(1 - x_6) + (1 - x_4)x_3x_6 \in H_3.$$

*In the above decomposition, $f_{1234}$ and $f_{456}$ are the polynomials corresponding to the two cliques $\{1, 2, 3, 4\}$ and $\{4, 5, 6\}$ (obtained by adding the chords 34 and 46 to $G_2$), and the polynomial $f'_{4,36}$ permits to cancel the quadratic terms $x_3x_4$ and $x_4x_6$ corresponding to the added edges 34 and 46 and to add the quadratic term $x_3x_6$.*

*This construction extends easily to an arbitrary $k \geq 3$, showing $\mathrm{rk_H}(G_k) \leq 4$. For example, $\alpha(G_3) - p_{G_3} = f_{1234} + f_{4567} + f_{789} + f'_{4,36} + f'_{7,69} \in H_4$.*

*Observe that the upper bound from Corollary 6.24 is not strong enough to show this. Indeed the defect of the inequality $\sum_{i \in V(G_k)} x_i \leq \alpha(G_k) = k$ is equal to $2(\alpha^*(G_k) - \alpha(G_k)) = k$, since $\alpha(G_k) = k$ and $\alpha^*(G_k) = 3k/2$ (this follows from the fact $\sum_{i \in V(G_k)} x_i \leq \alpha(G_k)$ defines a facet of $\mathrm{ST}(G_k)$, shown in [62, Lemma 32 and Theorem 34], so that $\alpha^*(G_k) = 3k/2$ by Lemma 2.10 of [65]). Thus Corollary 6.24 permits only to conclude that $\mathrm{rk_H}(G_k) \leq k + 2$.*

### 6.2.4 Graph operations

In this section, we investigate the behavior of the Handelman rank under some graph operations like node or edge deletion, edge contraction, and taking clique sums. For simplicity, we only consider unweighted graphs, while some of the results can easily be extended to the weighted case.

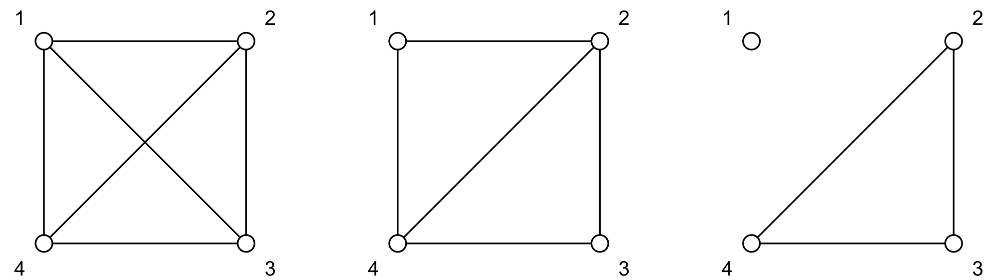


Figure 6.4: Example: edge deletion

### Operations on edges and nodes

An interesting observation is that the Handelman rank is not monotone under edge deletion. As an illustration, look at the three graphs in Figure 6.4. The first graph is the complete graph $K_4$, with $\mathrm{rk_H}(K_4) = 4$. If we delete one edge (say edge 13), we obtain the second graph $G$ with rank $\mathrm{rk_H}(G) = 2$. However, if we additionally delete the edges 12 and 14, then the third graph $G' = K_4 \setminus \{12, 13, 14\}$ has $\mathrm{rk_H}(G') = 3$, since it is the clique 0-sum of a node and a clique of size 3. (See Lemma 6.36 below.) On the other hand, if we delete an edge whose deletion increases the stability number (a so-called *critical edge*), then the Handelman rank does not increase.

**Lemma 6.33.** *Let $e$ be an edge of $G$ such that $\alpha(G \setminus e) = \alpha(G) + 1$. Then, one has*

$$\mathrm{rk_H}(G \setminus e) \leq \mathrm{rk_H}(G).$$

**Proof.** Say $e$ is the edge 12. Then, $\alpha(G \setminus e) - p_{G\setminus e} = \alpha(G) - p_G + 1 - x_1x_2$. As $1 - x_1x_2 = 1 - x_2 + x_2(1 - x_1) \in H_2$, this implies that $\mathrm{rk_H}(G \setminus e) \leq \mathrm{rk_H}(G)$. □

The Handelman rank is not monotone under edge contraction either. For instance, the graph $G$ in Figure 6.1 has $\mathrm{rk_H}(G) = 2$. If we contract the edge 23, we get the new graph $G'$ is a triangle with $\mathrm{rk_H}(G') = 3$. If we contract one more edge 12, the resulting graph $G''$ is an edge with $\mathrm{rk_H}(G'') = 2$. Analogously, deleting a node can either increase, decrease or not affect the Handelman rank. We group several properties about the behavior of the Handelman rank under node deletion.

**Lemma 6.34.** *Let $G = (V, E)$ be a graph and $j \in V$.*

*(i) If $\alpha(G - j) = \alpha(G)$, then* $\mathrm{rk_H}(G - j) \leq \mathrm{rk_H}(G)$.

*(ii) If $\alpha(G - j) = \alpha(G) - 1$, then* $\mathrm{rk_H}(G) \leq \mathrm{rk_H}(G - j)$.

*(iii) If $j$ is adjacent to all other nodes of $G$, then* $\mathrm{rk_H}(G) \leq \mathrm{rk_H}(G-j)+1$.

*(iv) If $j$ is an isolated node, then* $\mathrm{rk_H}(G) = \mathrm{rk_H}(G-j)$.

**Proof.** (i) We use relation (6.25) applied to the polynomial $f_G$ (and node $j$). As before $\underline{x}$ consists of all variables except $x_j$, so that $x = (\underline{x}, x_j)$. As $\alpha(G-j) = \alpha(G)$, we have $f_{G-j}(\underline{x}) = f_G(\underline{x}, 0) \in H_{\mathrm{rk_H}(G)}$, which implies $\mathrm{rk_H}(G-j) \leq \mathrm{rk_H}(G)$.

(ii) If $\alpha(G-j) = \alpha(G)-1$, then $f_G = f_{G-j} + (1-x_j) + \sum_{i:ij\in E} w_{ij}x_ix_j \in H_{\mathrm{rk_H}(G-j)}$. Hence, $\mathrm{rk_H}(G) \leq \mathrm{rk_H}(G-j)$.

(iii) Assume that $j$ is adjacent to all other nodes of $G$. If $G-j$ has no edge then $G$ is bipartite and thus $\mathrm{rk_H}(G) = 2 = \mathrm{rk_H}(G-j)+1$. Assume now that $G-j$ has an edge so that $\mathrm{rk_H}(G-j) \geq 2$. Using Lemma 6.21, we deduce that $\mathrm{rk_H}(G) \leq \mathrm{rk_H}(G-j)+1$.

(iv) $G$ is the clique 0-sum of $G-j$ and the single node $j$, and we can apply Lemma 6.36 below. □

**Remark 6.35.** *In Lemma 6.34 (ii), the gap* $\mathrm{rk_H}(G-j) - \mathrm{rk_H}(G)$ *can be arbitrarily large. To see this consider the graph $G$ obtained by taking the clique $t$-sum of $K_{2t}$ and $K_{t+1}$ along a common $K_t$. Let $j$ be the node of $K_{t+1}$ which does not belong to the common clique $K_t$. If we delete node $j$, then $G-j = K_{2t}$ has* $\mathrm{rk_H}(G-j) = 2t$. *On the other hand,* $\mathrm{rk_H}(G) \leq t+1$, *since $\alpha(G) = 2 = \rho_{t+1}(G)$ as $V(G)$ can be covered by two cliques of size at most $t+1$. Thus* $\mathrm{rk_H}(G-j) - \mathrm{rk_H}(G) \geq 2t-(t+1) = t-1$.

### Clique sums

Suppose $G = (V, E)$ is the clique $t$-sum of two graphs $G_1$ and $G_2$. We now study the Handelman rank of $G$, whose value needs technical case checking, depending on the values of the stability numbers of $G$, $G_1$, $G_2$ and of some subgraphs.

**Lemma 6.36.** *Suppose $G$ is the clique $t$-sum of $G_1$ and $G_2$ along a common $t$-clique $C_0$ and let $H_i = G_i \setminus C_0$ for $i = 1, 2$. The following holds.*

*(i) If $\alpha(G) = \alpha(G_1) + \alpha(G_2)$, then*

$$\mathrm{rk_H}(G) \leq \min\{\max\{\mathrm{rk_H}(G_1), \mathrm{rk_H}(H_2)\}, \max\{\mathrm{rk_H}(H_1), \mathrm{rk_H}(G_2))\}\}.$$

*Moreover,* $\mathrm{rk_H}(G) \leq \max\{\mathrm{rk_H}(G_1), \mathrm{rk_H}(G_2)\}$ *if $t \leq 3$.*

*(ii) Assume $\alpha(G) = \alpha(G_1) + \alpha(G_2) - 1$. Then $\alpha(G_k) = \alpha(H_k) + 1$ for (say) $k = 1$ and* $\mathrm{rk_H}(G) \leq \max\{\mathrm{rk_H}(H_1), \mathrm{rk_H}(G_2)\}$.

(iii) *Assume* $\alpha(G) = \alpha(G_1) + \alpha(G_2) - 2$. *For* $k \in \{1, 2\}$ *let* $C_k$ *denote the set of nodes of* $C_0$ *which belong to at least one maximum stable set of* $G_k$. *Set* $H_1' = G_1 \backslash C_1$ *and* $H_2' = G \setminus H_1' = G_2 \setminus (C_0 \setminus C_1)$. *Then* $\alpha(H_k') = \alpha(G_k) - 1$ *for* $k = 1, 2$, *and* $\mathrm{rk_H}(G) \leq \max\{\mathrm{rk_H}(H_1'), \mathrm{rk_H}(H_2')\}$.

**Proof.** In what follows, for subsets $A, B \subseteq V$, $E(A, B)$ denotes the set of edges $ij$ with $i \in A$ and $j \in B$, and $E(A)$ the set of edges contained in $A$. We also set $VG$ for $V(G)$.

(i) We use the identities

$$f_G = f_{G_1} + f_{H_2} + (\alpha(G) - \alpha(G_1) - \alpha(H_2)) + \sum_{ij \in E(VG_1, VH_2)} x_i x_j,$$

$$f_G = f_{G_2} + f_{H_1} + (\alpha(G) - \alpha(G_2) - \alpha(H_1)) + \sum_{ij \in E(VG_2, VH_1)} x_i x_j.$$

As $\alpha(G) = \alpha(G_1) + \alpha(G_2)$, $\alpha(G) - \alpha(G_1) = \alpha(G_2) \geq \alpha(H_2)$ and $\alpha(G) - \alpha(G_2) = \alpha(G_1) \geq \alpha(H_1)$, implying

$$\mathrm{rk_H}(G) \leq \min\{\max\{\mathrm{rk_H}(G_1), \mathrm{rk_H}(H_2)\}, \max\{\mathrm{rk_H}(G_2), \mathrm{rk_H}(H_1)\}\}.$$

For the second statement, we use the identity

$$f_G = f_{G_1} + f_{G_2} + \sum_{i \in C_0} x_i - \sum_{ij \in E(C_0)} x_i x_j$$

combined with the fact that $\sum_{i \in C_0} x_i - \sum_{ij \in E(C_0)} x_i x_j \in H_2$ when $t = |C_0| \leq 3$. This is clear if $t \leq 1$ and follows from the identities $x_1 + x_2 - x_1 x_2 = x_1(1 - x_2) + x_2 \in H_2$ and $x_1 + x_2 + x_3 - x_1 x_2 - x_1 x_3 - x_2 x_3 = x_1(1 - x_2) + x_2(1 - x_3) + x_3(1 - x_1) \in H_2$ if $t = 2, 3$. From this follows that $\mathrm{rk_H}(G) \leq \max\{\mathrm{rk_H}(G_1), \mathrm{rk_H}(G_2)\}$.

(ii) As $\alpha(G) \neq \alpha(G_1) + \alpha(G_2)$, it follows that $\alpha(H_k) = \alpha(G_k) - 1$ for at least one index $k = 1, 2$. Say this holds for $k = 1$. Then we use the identities

$$f_G = f_{G_1} + f_{G_2} - 1 + \sum_{i \in C_0} x_i - \sum_{ij \in E(C_0)} x_i x_j,$$

and

$$f_{H_1} = f_{G_1} - 1 + \sum_{i \in C_0} x_i - \sum_{ij \in E(C_0)} x_i x_j - \sum_{ij \in E(C_0, VG_1 \backslash C_0)} x_i x_j.$$

This gives:

$$f_G = f_{H_1} + f_{G_2} + \sum_{ij \in E(C_0, VG_1 \setminus C_0)} x_i x_j,$$

which implies $\mathrm{rk}_{\mathrm{H}}(G) \le \max\{\mathrm{rk}_{\mathrm{H}}(H_1), \mathrm{rk}_{\mathrm{H}}(G_2)\}$.

(iii) By construction, $\alpha(H_1') = \alpha(G_1) - 1$. Moreover, as $\alpha(G) = \alpha(G_1) + \alpha(G_2) - 2$, it follows that $C_1 \cap C_2 = \emptyset$ and thus $\alpha(H_2') = \alpha(G_2) - 1$. We now use the identities

$$f_{H_1'} = f_{G_1} - 1 + \sum_{i \in C_1} x_i - \sum_{ij \in E(C_1) \cup E(C_1, VG_1 \setminus C_1)} x_i x_j,$$

$$f_{H_2'} = f_{G_2} - 1 + \sum_{i \in C_0 \setminus C_1} x_i - \sum_{ij \in E(C_0 \setminus C_1) \cup E(C_0 \setminus C_1, VG_2 \setminus (C_0 \setminus C_1))} x_i x_j,$$

and

$$f_G = f_{G_1} + f_{G_2} - 2 + \sum_{i \in C_0} x_i - \sum_{ij \in E(C_0)} x_i x_j.$$

Combining these relations, we obtain

$$f_G = f_{H_1'} + f_{H_2'} + \sum_{ij \in E(C_1, VG_1 \setminus C_1) \cup E(C_0 \setminus C_1, VG_2 \setminus C_0)} x_i x_j$$

which shows $\mathrm{rk}_{\mathrm{H}}(G) \le \max\{\mathrm{rk}_{\mathrm{H}}(H_1'), \mathrm{rk}_{\mathrm{H}}(H_2')\}$. □

In the special case when $G$ is a clique sum of two cliques, one can easily determine the the exact value of the Handelman rank of $G$.

**Lemma 6.37.** *Assume that $G$ is the clique $t$-sum of two cliques $K_{n_1}$ and $K_{n_2}$ with $n_1 \le n_2$. Then,* $\mathrm{rk}_{\mathrm{H}}(G) = \max\{\lceil \frac{n_1+n_2-t}{2} \rceil, n_2 - t\}$.

**Proof.** Obviously, $\alpha(G) = 2$. Define $n = |V(G)| = n_1 + n_2 - t$. Assume first that $n_2 - n_1 \le t$. Then $V(G)$ can be covered by two cliques of sizes $\lceil \frac{n}{2} \rceil$ and $\lfloor \frac{n}{2} \rfloor$ and thus $\mathrm{rk}_{\mathrm{H}}(G) \le \lceil \frac{n}{2} \rceil$. In addition, by (6.23), $\mathrm{rk}_{\mathrm{H}}(G) \ge \frac{n}{\alpha(G)} = \frac{n}{2}$. Hence we obtain $\mathrm{rk}_{\mathrm{H}}(G) = \lceil \frac{n}{2} \rceil = \max\{\lceil \frac{n}{2} \rceil, n_2 - t\}$.

Assume now that $n_2 - n_1 > t$. Then $G$ can be covered by two cliques of sizes $n_1$ and $n_2 - t$, which implies $\mathrm{rk}_{\mathrm{H}}(G) \le n_2 - t$. On the other hand, by applying Lemma 6.34 (i) to all nodes $i$ in the common $t$-clique, together with Lemma 6.36, we obtain the reverse inequality $\mathrm{rk}_{\mathrm{H}}(G) \ge \max\{\mathrm{rk}_{\mathrm{H}}(K_{n_2-t}), \mathrm{rk}_{\mathrm{H}}(K_{n_1-t})\} = n_2 - t$. □

## 6.3 Links to other hierarchies

Several other hierarchies have been considered in the literature for general 0-1 optimization problems, which can also be applied to the maximum stable set problem in (6.11). In this section we review the hierarchies proposed by Sherali and Adams [90], Lovász and Schrijver [65], Lasserre [49], and De Klerk and Pasechnik [25], respectively. More over, we briefly indicate how they relate to the Handelman hierarchy considered in this chapter, based on optimization on the hypercube.

### 6.3.1 Sherali-Adams and Lasserre hierarchies

Consider the following 0-1 polynomial optimization problem:

$$\max \quad f(x) \quad \text{s.t.} \quad x \in \mathbf{K} \cap \{0,1\}^n, \tag{6.28}$$

which is obtained by adding the integrality constraint $x \in \{0,1\}^n$ to the problem of maximizing the polynomial $f$ over the set

$$\mathbf{K} = \{x \in \mathbb{R}^n : g_1(x) \geq 0, \ldots, g_m(x) \geq 0\}.$$

Recall that $\mathcal{I}$ denotes the ideal generated by $x_i - x_i^2$ for $i \in [n]$ and that the Handelman set $H_r$ is defined in (6.3). Sherali and Adams [90] introduce the following bounds for (6.28):

$$f_{\text{sa}}^{(r)} = \inf \left\{ \lambda : \lambda - f \in H_r + \sum_{j=1}^{m} g_j H_{r-\deg(g_j)} + \mathcal{I} \right\}. \tag{6.29}$$

The above program is in fact the dual of the linear program usually used to define the Sherali-Adams bounds. For details we refer, e.g., to [90, 50, 55].

When applying the Sherali-Adams construction to the maximum stable set problem for the instance $(G, w)$, the starting point is to formulate $\alpha(G, w)$ as the problem of maximizing the linear polynomial $p(x) = w^T x = \sum_{i \in [n]} w_i x_i$ over $\mathbf{K} \cap \{0,1\}^n$, where $\mathbf{K} = \text{FR}(G)$ is the fractional stable set polytope, so that the corresponding bound from (6.29) reads

$$p_{\text{sa}}^{(r)}(G, w) = \inf \left\{ \lambda : \lambda - w^T x \in H_r + \sum_{ij \in E} (1 - x_i - x_j) H_{r-1} + \mathcal{I} \right\}. \tag{6.30}$$

For $r \geq 2$, let $\langle x_i x_j : ij \in E\rangle_r$ denote the truncated ideal consisting of all polynomials $\sum_{ij\in E} u_{ij}x_ix_j$ where $u_{ij} \in \mathbb{R}[x]$ has degree at most $r-2$. One can formulate the following variation of the bound (6.30):

$$\mathrm{sa}^{(r)}(G,w) = \inf\{\lambda : \lambda - w^T x \in H_r + \langle x_ix_j : ij \in E\rangle_r + \mathcal{I}\},$$

which satisfies

$$\mathrm{sa}^{(r+1)}(G,w) \leq p_{\mathrm{sa}}^{(r)}(G,w) \leq \mathrm{sa}^{(r)}(G,w).$$

To see it use, for any edge $ij \in E$, the identities $1 - x_i - x_j = (1-x_i)(1-x_j) - x_ix_j$ and $-x_ix_j = x_i(1-x_i-x_j) + x_i(x_i-1)$.

Comparing with the hypercube based Handelman bound (6.12), we see that

$$\mathrm{sa}^{(r)}(G,w) \leq \bar{p}_{\mathrm{han}}^{(r)}(G,w),$$

since $\lambda - p_{G,w} = \lambda - w^T x + \sum_{ij\in E} w_{ij}x_ix_j \in H_r$ implies $\lambda - w^T x \in H_r + \langle x_ix_j : ij \in E\rangle_r$.

We now recall the following semidefinite programming bound of Lasserre [49]:

$$\mathrm{las}^{(r)}(G,w) = \inf\{\lambda : \lambda - w^T x \in \Sigma_r + \langle x_ix_j : ij \in E\rangle_r + \mathcal{I}\},$$

where $\Sigma_r$ is the set of sums of squares of polynomials of degree at most $2r$.

As is well known,

$$\mathrm{las}^{(r)}(G,w) \leq \mathrm{sa}^{(r)}(G,w).$$

This can easily be seen by noting that, for any set $T$ with $|T| = r$, we have

$$x^I(1-x)^{T\setminus I} = \underbrace{\prod_{i\in I} x_i^2 \prod_{j\in T\setminus I}(1-x_j)^2}_{\in\Sigma_r} + \underbrace{\left(\prod_{i\in I} x_i \prod_{j\in T\setminus I}(1-x_j) - \prod_{i\in I} x_i^2 \prod_{j\in T\setminus I}(1-x_j)^2\right)}_{\in\mathcal{I}},$$

where the second term belongs to $\mathcal{I}$ in view of Lemma 6.3. Summarizing, we have

$$\alpha(G,w) \leq \mathrm{las}^{(r)}(G,w) \leq \mathrm{sa}^{(r)}(G,w) \leq \bar{p}_{\mathrm{han}}^{(r)}(G,w).$$

Hence, the Sherali-Adams and Lasserre bounds are at least as strong as the Handelman bound at any given order $r$. Thus, the error estimate for the Handelman bound $\bar{p}_{\mathrm{han}}^{(r)}(G,w)$ in Theorem 6.7 also holds for the bounds $\mathrm{las}^{(r)}(G,w)$ and $\mathrm{sa}^{(r)}(G,w)$.

On the other hand, the Sherali-Adams and Lasserre bounds are more expensive to compute. Indeed the Sherali-Adams bound is linear but its definition involves more terms, and the Lasserre bound is based on semidefinite programming which is computationally more demanding than linear programming. For more results about the comparison between Sherali-Adams and Lasserre hierarchies, see, e.g., [50, 55].

### 6.3.2 Lovász-Schrijver hierarchy

Given a polytope $\mathbf{K} \subseteq [0,1]^n$, Lovász and Schrijver [65] build a hierarchy of polytopes nested between $\mathbf{K}$ and the convex hull of $\mathbf{K} \cap \{0,1\}^n$ that finds it after $n$ steps. When applied to the maximum stable set problem, one starts with the fractional stable set polytope $\mathbf{K} = \mathrm{FR}(G)$. For convenience set $\widehat{V} = V \cup \{0\}$ (where 0 is an additional element not belonging to $V$) and define the cone

$$\mathcal{C}(G) = \left\{ \lambda \begin{pmatrix} 1 \\ x \end{pmatrix} : x \in \mathrm{FR}(G), \lambda \geq 0 \right\} \subseteq \mathbb{R}^{\widehat{V}}.$$

Define the following set of symmetric matrices indexed by $\widehat{V}$:

$$\mathcal{M}(G) = \{Y \in \mathcal{S}_{\widehat{V}} : Y_{ii} = Y_{0i} \ \forall i \in V, \ Y\mathbf{e}_i, Y(\mathbf{e}_0 - \mathbf{e}_i) \in \mathcal{C}(G) \ \forall i \in V\}$$

and the corresponding subset of $\mathbb{R}^{|V|}$:

$$N(\mathrm{FR}(G)) = \left\{ x \in \mathbb{R}^{|V|} : \begin{pmatrix} 1 \\ x \end{pmatrix} = Y\mathbf{e}_0 \ \text{ for some } \ Y \in \mathcal{M}(G) \right\}.$$

For $r \geq 2$, define the $r$-th iterate $N^r(\mathrm{FR}(G)) = N(N^{r-1}(\mathrm{FR}(G)))$, setting $N^1(\mathrm{FR}(G)) = N(\mathrm{FR}(G))$. It is shown in [65] that

$$\mathrm{ST}(G) \subseteq \ldots \subseteq N^r(\mathrm{FR}(G)) \subseteq N^{r-1}(\mathrm{FR}(G)) \subseteq \ldots \subseteq N(\mathrm{FR}(G)) \subseteq \mathrm{FR}(G),$$

with equality $\mathrm{ST}(G) = N^n(\mathrm{FR}(G))$. By maximizing the linear function $w^T x$ over $N^r(\mathrm{FR}(G))$ we get the bound $\mathrm{ls}^{(r)}(G,w)$ which satisfies $p_{\mathrm{sa}}^{(r+1)}(G,w) \leq \mathrm{ls}^{(r)}(G,w)$ for $r \geq 1$ (see [65, 55]).

For any $w \in \mathbb{R}_+^{|V|}$, the corresponding inequality $w^T x \leq \alpha(G,w)$ is valid for $\mathrm{ST}(G)$. Following [65], its $N$*-index*, denoted here as $\mathrm{rk}_{\mathrm{LS}}(G,w)$, is the smallest integer $r$ for which the inequality $w^T x \leq \alpha(G,w)$ is valid for $N^r(\mathrm{FR}(G))$ or, equivalently, $\alpha(G,w) = \mathrm{ls}^{(r)}(G,w)$. The following bounds are shown in [65] for the $N$-index:

$$\frac{\sum_{i=1}^n w_i}{\alpha(G,w)} - 2 \leq \mathrm{rk}_{\mathrm{LS}}(G,w) \leq \mathrm{defect}(G,w), \ \mathrm{rk}_{\mathrm{LS}}(G,w) \leq |V(G)| - \alpha(G) - 1,$$

where $\mathrm{defect}(G,w)$ is as defined in (6.26). Note the analogy with the bounds (6.23), (6.24) and (6.27) for the Handelman rank. There is a shift of 2 between the two hierarchies which can be explained from the fact that the Lovász-Schrijver construction starts from the fractional stable set polytope which already takes the edges into account, so that $\mathrm{ls}^{(0)}(G,w) = \alpha^*(G,w) = \bar{p}_{\mathrm{han}}^{(2)}(G,w)$. We also observe this shift by 2, e.g., in the results for perfect graphs and for odd circuits and wheels. It seems moreover that the Handelman bound and the bound obtained by using the $N$-operator are closely related. We did some computational tests for the graphs $K_4$, $W_5$ and $G_k$ ($k = 2, 3, 4, 5$) with different weight functions; in all cases we observe that both bounds coincide, i.e., $\mathrm{ls}^{(1)}(G, w) = \bar{p}^{(3)}_{\mathrm{han}}(G, w)$ holds. Understanding the exact link between the two hierarchies of Handelman and of Lovász-Schrijver is an interesting open question.

### 6.3.3 De Klerk and Pasechnik LP hierarchy

Given a graph $G = (V, E)$ with adjacency matrix $A$, De Klerk and Pasechnik [25] formulate its stability number via the following copositive program:

$$\alpha(G) = \min\{\lambda : \lambda(I + A) - \mathbf{e}\mathbf{e}^T \in \mathcal{C}_n\},$$

which is based on the Motzkin-Straus formulation:

$$\frac{1}{\alpha(G)} = \min_{x \in \Delta_{|V|}} x^T(I + A)x, \tag{6.31}$$

where $I$ denotes the identity matrix. As problem (6.31) is the problem of minimizing the quadratic polynomial $q(x) = x^T(I + A)x$ over the standard simplex $\Delta_{|V|}$, one can apply Lemma 1.4 and define, for any $r \geq 2$, the corresponding Handelman bound

$$q^{(r)}_{\mathrm{han}} = \max\{\lambda : (q - \lambda\sigma^2)\sigma^{r-2} \in \mathbb{R}_+[x]\},$$

where $\sigma = \sum_{i=1}^n x_i$. It turns out that it can be computed explicitly since it is directly related to the following bound introduced in [25]:

$$\zeta^{(r)}(G) = \min\{\mu : (\mu q - \sigma^2)\sigma^r \in \mathbb{R}_+[x]\}$$

for any $r \geq 0$. Indeed it follows from the definitions that

$$\zeta^{(r)} q^{(r+2)}_{\mathrm{han}} = 1 \quad \text{for } r \geq 0.$$

De Klerk and Pasechnik [25] show that

$$\zeta^{(0)}(G) \geq \zeta^{(1)}(G) \geq \cdots \geq \lfloor \zeta^{(r)}(G) \rfloor = \alpha(G)$$

for $r \geq \alpha(G)^2 - 1$. Moreover, Peña, Vera and Zuluaga [81] give the following closed-form expression for the parameter $\zeta^{(r)}(G)$:

$$\zeta^{(r)}(G) = \frac{\binom{r+2}{2}}{\binom{u}{2}\alpha(G) + uv}, \quad \text{where } r + 2 = u\alpha(G) + v \text{ with } u, v \in \mathbb{N} \text{ and } v < \alpha(G).$$

From this we see that $\zeta^{(r)}(G) = \infty$ if $r \leq \alpha(G) - 2$ and $\zeta^{(r)}(G) = \alpha(G) + 1$ if $r = \alpha(G)^2 - 2$. Moreover, $\alpha(G) \leq \zeta^{(r)}(G) < \alpha(G) + 1$ for any $r \geq \alpha(G)^2 - 1$, with a strict inequality $\alpha(G) < \zeta^{(r)}(G)$ if $G$ is not a complete graph. Hence, in contrast to the LP bounds based on the Handelman, Sherali-Adams and Lovász-Schrijver constructions (which are exact at order $n$), the LP copositive-based bound is never exact (except for the complete graph), one needs to round it in order to obtain the stability number.

From the above discussion it follows that the LP copositive rank $\mathrm{rk_{KP}}(G)$, which we define as the smallest integer $r$ such that $\lfloor \zeta^{(r)}(G) \rfloor = \alpha(G)$, can be determined exactly: $\mathrm{rk_{KP}}(G) = \alpha(G)^2 - 1$ for any graph $G$. We now observe that it cannot be compared with the (hypercube based) Handelman rank $\mathrm{rk_H}(G)$. Indeed, for the complete graph $G = K_n$, we have $\mathrm{rk_{KP}}(K_n) = 0$ while $\mathrm{rk_H}(K_n) = n$. On the other hand, the graph $K_{1,n}$ has $\mathrm{rk_{KP}}(K_{1,n}) = n^2 - 1$ and $\mathrm{rk_H}(K_{1,n}) = 2$. As another example, for the graph $G_k$ from Example 6.32, $\mathrm{rk_{KP}}(G_k) = k^2 - 1$ while $\mathrm{rk_H}(G_k) \leq 4$. Hence the ranks of the two hierarchies are not comparable. These examples also show that the ranks of the Lovász-Schrijver and of the LP copositive hierarchies are not comparable, since $\mathrm{rk_{LS}}(K_n) = n - 2$ and $\mathrm{rk_{LS}}(K_{1,n}) = 0$.

## 6.4 The Handelman hierarchy for the maximum cut problem

We now conclude with some observations clarifying how the Handelman hierarchy applies to the maximum cut problem. Given a graph $G = (V, E)$ with edge weights $w \in \mathbb{R}^{|E|}$, the max-cut problem asks to find a partition $(V_1, V_2)$ of the node set $V$ so that the total weight of the edges cut by the partition is maximized; it is NP-hard, already in the unweighted case [41]. As observed in [77] the maximum weight of a cut in $G$ can be computed via the following problem (as in (1.4)):

$$\mathrm{mc}(G, w) = \max_{x \in [0,1]^{|V|}} \sum_{i \in V} d_i x_i - 2 \sum_{ij \in E} w_{ij} x_i x_j,$$

setting $d_i = \sum_{j \in V : ij \in E} w_{ij}$. As the polynomial $f(x) = \sum_{i \in V} d_i x_i - 2 \sum_{ij \in E} w_{ij} x_i x_j$ is square-free the Handelman bound of order $t$ can be formulated as

$$\min\{\lambda : \lambda - f \in H_r\}.$$

We show below that it can be equivalently reformulated in a more explicit way in terms of suitable valid inequalities for the cut polytope. We need some definitions.

The cut polytope $\mathrm{CUT}_n$ is defined as the convex hull of the vectors $(v_iv_j)_{1\le i<j\le n}$ for all $v \in \{\pm 1\}^n$. So $\mathrm{CUT}_n$ is a polytope in the space $\mathbb{R}^{\binom{n}{2}}$ indexed by the edge set of the complete graph $K_n$. Given an integer $r \geq 2$, among all the inequalities that are valid for $\mathrm{CUT}_n$, we consider only those that are supported by at most $r$ points of $[n]$ and we let $P_n^{(r)}$ denote the polytope in $\mathbb{R}^{\binom{n}{2}}$ defined by all these selected inequalities. Clearly, $\mathrm{CUT}_n \subseteq P_n^{(r)}$. Moreover, for $n \neq 4$, equality $\mathrm{CUT}_n = P_n^{(r)}$ holds if and only if $r = n$ (since $\mathrm{CUT}_n$ has some facet defining inequalities supported by $n$ points). The case $n = 4$ is an exception since $\mathrm{CUT}_4 = P_4^{(3)}$.

**Proposition 6.38.** *Let $r \geq 2$ and, given an edge weighted graph $(G, w)$, consider the above mentioned polynomial $f = \sum_{i\in V} d_ix_i - 2\sum_{ij\in E} w_{ij}x_ix_j$. The following equality holds:*

$$\min\{\lambda : \lambda - f \in H_r\} = \max_{y\in P_n^{(r)}} \sum_{ij\in E} w_{ij}(1 - y_{ij})/2.$$

**Proof.** It is convenient to use $\pm 1$ valued variables $z$ instead of the 0-1 valued variables $x$. So we set $z_i = 1 - 2x_i$ for $i \in [n]$. Then $f(x) = q(z)$, after defining the polynomial $q(z) = \sum_{ij\in E} w_{ij}(1 - z_iz_j)/2$. Moreover define the $\pm 1$ analogue of the Handelman set $H_r$ from (6.3):

$$\overline{H}_r = \{ \sum_{T\subseteq[n]:|T|=r} \sum_{I\subseteq T} c_{I,T}(1 - z)^I (1 + z)^{T\setminus I} : c_{I,T} \geq 0\}.$$

Furthermore let $\overline{\mathcal{I}}$ denote the ideal in the polynomial ring $\mathbb{R}[z]$ generated by $z_i^2 - 1$ for $i \in [n]$, and let $\overline{\mathcal{I}}_r$ denote its truncation at degree $r$. One can easily verify that $\lambda - f \in H_r$ if and only if $\lambda - q \in \overline{H}_r$ which, in turn, is equivalent to $\lambda - q \in \overline{H}_r + \overline{\mathcal{I}}_r$. Therefore we have

$$\min\{\lambda : \lambda - f \in H_r\} = \min\{\lambda : \lambda - q \in \overline{H}_r + \overline{\mathcal{I}}_r\}.$$

Now we apply LP duality and obtain that the last program is equal to

$$\max_{\mathcal{L}}\{\mathcal{L}(q) : \mathcal{L}(1) = 1,\ \mathcal{L}(f) \geq 0\ \forall f \in \overline{H}_r,\ \mathcal{L}(f) = 0\ \forall f \in \overline{\mathcal{I}}_r\},$$

where the maximum is taken over all linear functionals $\mathcal{L} : \mathbb{R}[z]_r \to \mathbb{R}$. Finally, we use the fact that this maximization program is equal to the maximum of

$$\sum_{ij\in E} \frac{w_{ij}(1 - y_{ij})}{2}$$

taken over all $y \in P_n^{(r)}$, which is shown in [55]. This concludes the proof. □

For instance, for $r = 2$, $P_n^{(2)} = [-1,1]^{\binom{n}{2}}$ (since $-1 \leq y_{ij} \leq 1$ are the only inequalities on two points valid for $\mathrm{CUT}_n$). Hence, by Proposition 6.38, the Handelman bound of order 2 is equal to $\sum_{ij \in E} |w_{ij}|$, as shown in [77] for the case $w \geq 0$. For $r = 3$, $P_n^{(3)}$ is defined by the triangle inequalities $y_{ij} + y_{ik} + y_{jk} \geq -1$ and $y_{ij} - y_{ik} - y_{jk} \geq -1$ for all $i, j, k \in [n]$. Therefore, for an edge weighted graph $G$ where $G$ has no $K_5$ minor, we find that the Handelman bound of order 3 is exact and returns the value of the maximum cut (since the triangle inequalities suffice to describe the cut polytope of $G$, after taking projections; see, e.g., [4] or [56, Section 3.6]). In particular, the Handelman rank is at most 3 for a weighted odd circuit, which answers an open question of [78] (which shows the result in the unweighted case). As a final observation, we find that the rank of the Handelman hierarchy for the maximum cut problem in $K_n$ is equal to $n$ for any $n \neq 4$ (which was shown in [77] for $n$ odd).

# Appendix A

# Stirling numbers of the second kind

In this appendix, we review the necessary background material on Striling numbers of the second kind. Recall that, for integers $n, k \in \mathbb{N}$, the Stirling number of the second kind $S(n, k)$ counts the number of ways of partitioning a set of $n$ objects into $k$ nonempty subsets. Thus $S(n, k) = 0$ if $k > n$, $S(n, 0) = 0$ if $n \geq 1$, and $S(0, 0) = 1$ by convention.

We first review the following recursive relation for Stirling numbers of the second kind.

**Lemma A.1.** *For any integers $\beta \geq 0$ and $\alpha \geq 1$, one has*

$$S(\beta + 1, \alpha) = S(\beta, \alpha - 1) + \alpha S(\beta, \alpha). \tag{A.1}$$

**Proof.** This well known fact can be easily checked as follows. By definition, $S(\beta + 1, \alpha)$ counts the number of ways of partitioning the set $\{1, \ldots, \beta, \beta + 1\}$ into $\alpha$ nonempty subsets. Considering the last element $\beta + 1$, one can either put it in a singleton subset (so that there are $S(\beta, \alpha - 1)$ such partitions), or partition $\{1, \ldots, \beta\}$ into $\alpha$ nonempty subsets and then assign the last element $\beta + 1$ to one of them (so that there are $\alpha S(\beta, \alpha)$ such partitions). This shows the result. $\square$

We next recall the following expression for Stirling numbers of the second kind. We provide a full proof, since we could not find this result in the literature.

**Lemma A.2.** *Given $\alpha \in I(n, k)$ and $d > k$, one has*

$$S(d, k) = \frac{\alpha!}{k!} \sum_{\beta \in I(n,d)} \frac{d!}{\beta!} \prod_{i=1}^{n} S(\beta_i, \alpha_i). \tag{A.2}$$

**Proof.** For integers $d, k \geq 0$, let $S_{d,k}$ denote the number of surjective maps from a $d$-elements set to a $k$-elements set. It is not difficult to see the following relation between $S_{d,k}$ and $S(d,k)$:

$$S_{d,k} = k!S(d,k).$$

Indeed, let $B = [d]$ and $A = [k]$. In order to choose a surjective map $f$ from $B$ to $A$ one needs to select the pre-image $B_i = f^{-1}(i) \subseteq B$ for each element $i \in [k]$. So to define a surjective map $f$, one first selects a partition of $B$ into $k$ non-empty subsets $B_1, \ldots, B_k$, which can be done in $S(d,k)$ ways. As any permutation of the $B_i's$ gives rise to a distinct surjective map, there are $k!S(d,k)$ surjective maps from $[d]$ to $[k]$.

Now, the identity (A.2) about the Stirling numbers $S(d,k)$ can be equivalently reformulated as the following identity about the numbers $S_{d,k}$: For any $\alpha \in I(n,k)$,

$$S_{d,k} = \sum_{\beta \in I(n,d)} \frac{d!}{\beta!} \prod_{i=1}^{n} S_{\beta_i,\alpha_i}.$$

Again set $B = [d]$ and $A = [k]$. Say, $\alpha$ has $p$ non-zero coordinates, i.e., $\alpha_1, \ldots, \alpha_p \geq 1$ and $\alpha_1 + \ldots + \alpha_p = k$. Fix a partition of $A = [k]$ into $p$ subsets $A_1, \ldots, A_p$ where $|A_i| = \alpha_i$ for $i \in [p]$. Then, a surjection $f$ from $B$ to $A$ defines a surjection from $B_i = f^{-1}(A_i)$ to $A_i$ for each $i \in [p]$. Setting $\beta_i = |B_i|$, we have $\beta_1 + \ldots + \beta_p = d$ since the $B_i's$ partition $B$. Hence one can count the number of surjections from $B$ to $A$ as follows.

First, select $\beta_1, \ldots, \beta_p \geq 1$ such that $\beta_1 + \ldots + \beta_p = d$. Then split the $d$ elements of $B$ into an ordered sequence of $p$ disjoint subsets $B_1, \ldots, B_p$ where $|B_i| = \beta_i$ for $i \in [p]$; there are $\frac{d!}{\beta!}$ ways of doing so. Once $B_1, \ldots, B_p$ are selected, there are $S_{\beta_i,\alpha_i}$ possible surjections from $B_i$ to $A_i$ for each $i \in [p]$ and thus a total of $\prod_{i=1}^{p} S_{\beta_i,\alpha_i}$ possibilities. Therefore, we get that the total number of surjections from $B$ to $A$ is equal to $\sum_{\beta \in I(p,d)} \frac{d!}{\beta!} \prod_{i=1}^{p} S_{\beta_i,\alpha_i}$, which shows the result. $\square$

Finally we recall the following result, implied by [44, relation (3.2)], and we only sketch the proof.

**Lemma A.3.** *[44, relation (3.2)] For positive integers $d$ and $r \geq 1$, one has*

$$\sum_{k=1}^{d-1} r^{\underline{k}} S(d,k) = r^d - r^{\underline{d}}.$$

**Proof.** The proof is by induction on $d$, and using relation (A.1) for the induction step. $\square$

# Appendix B

# Proof of Theorem 2.18

In this appendix, we give a self-contained proof for Theorem 2.18, which provides an explicit description of the moments of the multinomial distribution in terms of the Stirling numbers of the second kind.

Fix $x = (x_1, \ldots, x_n) \in \Delta_n$ and assume $Y = (Y_1, \ldots, Y_n)$ has the multinomial distribution with parameters $r$, $n$ and $x_1, \ldots, x_n$ (where $x \in \Delta_n$) as in Section 2.2.1. Then, given $\alpha \in I(n, r)$, by (2.2), the probability of the event $Y = \alpha$ is equal to $\frac{r!}{\alpha!} x^\alpha$. Recall that, for $\beta \in \mathbb{N}^n$, the $\beta$-th moment of this multinomial distribution was given in (2.18):

$$m_{(n,r)}^{\beta} = \mathbb{E}[Y^\beta] = \sum_{\alpha \in I(n,r)} \alpha^\beta \frac{r!}{\alpha!} x^\alpha. \tag{B.1}$$

Our objective is to show the following reformulation of the $\beta$-th moment in terms of the Stirling numbers of the second kind:

$$m_{(n,r)}^{\beta} = \sum_{\alpha \in \mathbb{N}^n : \alpha \leq \beta} r^{\underline{|\alpha|}} x^\alpha \prod_{i=1}^{n} S(\beta_i, \alpha_i). \tag{B.2}$$

Our proof is elementary in the sense that we will obtain the moments of the multinomial distribution using its moment generating function. One of the ingredients which we will use is the fact that the identity (B.2) holds for the case $n = 2$ of the binomial distribution when $\beta \in \mathbb{N}^2$ is of the form $\beta = (\beta_1, 0)$. Namely, the following identity is shown in [44, Theorem 2.2 and relation (3.1)].

**Lemma B.1.** *[44] Given $\beta_1 \in \mathbb{N}$ and $x_1 \in \mathbb{R}$ such that $0 \leq x_1 \leq 1$, one has*

$$m_{(2,r)}^{(\beta_1,0)} = \sum_{\alpha_1=0}^{r} \alpha_1^{\beta_1} \binom{r}{\alpha_1} x_1^{\alpha_1} (1 - x_1)^{r-\alpha_1} = \sum_{\alpha_1=0}^{\beta_1} r^{\underline{\alpha_1}} x_1^{\alpha_1} S(\beta_1, \alpha_1).$$

This implies that the identity (B.2) holds for the moments of the multinomial distribution when the order $\beta$ has a single non-zero coordinate, i.e., $\beta$ is of the form $\beta = \beta_i \mathbf{e}_i$ with $\beta_i \in \mathbb{N}$.

**Corollary B.2.** *Given $\beta_i \in \mathbb{N}$ and $x \in \Delta_n$, one has*

$$m_{(n,r)}^{\beta_i \mathbf{e}_i} = \sum_{\alpha_i=0}^{\beta_i} r^{\underline{\alpha_i}} x_i^{\alpha_i} S(\beta_i, \alpha_i).$$

**Proof.** By (B.1), we have

$$\begin{aligned} m_{(n,r)}^{(\beta_i \mathbf{e}_i)} &= \sum_{\alpha \in I(n,r)} \alpha_i^{\beta_i} \frac{r!}{\alpha!} x^{\alpha} = \sum_{\alpha_i=0}^{r} {\alpha_i}^{\beta_i} \frac{r!}{\alpha_i!(r-\alpha_i)!} x_i^{\alpha_i} \left( \sum_{\overline{\alpha} \in I(n-1, r-\alpha_i)} \frac{(r-\alpha_i)!}{\overline{\alpha}!} x^{\overline{\alpha}} \right) \\ &= \sum_{\alpha_i=0}^{r} {\alpha_i}^{\beta_i} \binom{r}{\alpha_i} x_i^{\alpha_i} \left( \sum_{j \neq i} x_j \right)^{r-\alpha_i} = \sum_{\alpha_i=0}^{r} {\alpha_i}^{\beta_i} \binom{r}{\alpha_i} x_i^{\alpha_i} (1-x_i)^{r-\alpha_i}, \end{aligned}$$

which is equal to $\sum_{\alpha_i=0}^{\beta_i} r^{\underline{\alpha_i}} x_i^{\alpha_i} S(\beta_i, \alpha_i)$ by Lemma B.1. □

In order to determine the moments of the multinomial distribution we use its moment generating function

$$t \in \mathbb{R}^n \mapsto M_x^r(t) := \left( \sum_{i=1}^{n} x_i e^{t_i} \right)^r. \tag{B.3}$$

Then, for $\beta \in \mathbb{N}^n$, the $\beta$-th moment of the multinomial distribution is equal to the $\beta$-th derivative of the moment generating function evaluated at $t = 0$. Namely,

$$m_{(n,r)}^{\beta} = \left. \frac{\partial^{|\beta|} M_x^r(t)}{\partial t_1^{\beta_1} \cdots \partial t_n^{\beta_n}} \right|_{t=0}. \tag{B.4}$$

By Corollary B.2 we know that, for any $\beta_i \in \mathbb{N}$,

$$m_{(n,r)}^{(\beta_i \mathbf{e}_i)} = \left. \frac{\partial^{\beta_i} M_x^r(t)}{\partial t_i^{\beta_i}} \right|_{t=0} = \sum_{\alpha_i=0}^{\beta_i} S(\beta_i, \alpha_i) r^{\underline{\alpha_i}} x_i^{\alpha_i}. \tag{B.5}$$

Next we show an analogue of the above relation (B.5) for the evaluation of the $\beta_i \mathbf{e}_i$-th derivative of the moment generating function at any point $t \in \mathbb{R}^n$.

**Lemma B.3.** *For $x \in \Delta_n$, $\beta_i \in \mathbb{N}$ and $t \in \mathbb{R}^n$, one has*

$$\frac{\partial^{\beta_i} M_x^r(t)}{\partial t_i^{\beta_i}} = \sum_{\alpha_i=0}^{\beta_i} S(\beta_i, \alpha_i) r^{\underline{\alpha_i}} x_i^{\alpha_i} e^{\alpha_i t_i} M_x^{r-\alpha_i}(t).$$

**Proof.** To simplify notation we set $M^r = M_x^r(t)$. We show the result using induction on $\beta_i \geq 0$. The result holds clearly for $\beta_i = 0$ and also for $\beta_i = 1$ in which case we have

$$\frac{\partial M^r}{\partial t_i} = r x_i e^{t_i} M^{r-1}. \tag{B.6}$$

We now assume that the result holds for $\beta_i$ and we show that it also holds for $\beta_i + 1$. For this, using the induction assumption, we obtain

$$\begin{aligned}
\frac{\partial^{\beta_i+1} M^r}{\partial t_i^{\beta_i+1}} &= \frac{\partial}{\partial t_i} \frac{\partial^{\beta_i} M^r}{\partial t_i^{\beta_i}} = \frac{\partial}{\partial t_i} \left( \sum_{\alpha_i=0}^{\beta_i} S(\beta_i, \alpha_i) r^{\underline{\alpha_i}} x_i^{\alpha_i} e^{\alpha_i t_i} M^{r-\alpha_i} \right) \\
&= \sum_{\alpha_i=0}^{\beta_i} S(\beta_i, \alpha_i) r^{\underline{\alpha_i}} x_i^{\alpha_i} \frac{\partial}{\partial t_i} (e^{\alpha_i t_i} M^{r-\alpha_i}). \qquad \text{(B.7)}
\end{aligned}$$

Now, using (B.6), we can compute the last term as follows:

$$\frac{\partial}{\partial t_i}(e^{\alpha_i t_i} M^{r-\alpha_i}) = \alpha_i e^{\alpha_i t_i} M^{r-\alpha_i} + (r - \alpha_i) x_i e^{(\alpha_i+1)t_i} M^{r-\alpha_i-1}.$$

Plugging this into relation (B.7), we deduce

$$\begin{aligned}
&\frac{\partial^{\beta_i+1} M^r}{\partial t_i^{\beta_i+1}} \\
&= \sum_{\alpha_i=0}^{\beta_i} \alpha_i S(\beta_i, \alpha_i) r^{\underline{\alpha_i}} x_i^{\alpha_i} e^{\alpha_i t_i} M^{r-\alpha_i} + \sum_{\alpha_i=0}^{\beta_i} S(\beta_i, \alpha_i) r^{\underline{\alpha_i+1}} x_i^{\alpha_i+1} e^{(\alpha_i+1)t_i} M^{r-\alpha_i-1} \\
&= \sum_{\alpha_i=0}^{\beta_i} \alpha_i S(\beta_i, \alpha_i) r^{\underline{\alpha_i}} x_i^{\alpha_i} e^{\alpha_i t_i} M^{r-\alpha_i} + \sum_{\alpha_i'=1}^{\beta_i+1} S(\beta_i, \alpha_i' - 1) r^{\underline{\alpha_i'}} x_i^{\alpha_i'} e^{\alpha_i' t_i} M^{r-\alpha_i'} \\
&= \sum_{\alpha_i=0}^{\beta_i} (\underbrace{\alpha_i S(\beta_i, \alpha_i) + S(\beta_i, \alpha_i - 1)}_{=S(\beta_i+1, \alpha_i) \text{ by (A.1)}}) r^{\underline{\alpha_i}} x_i^{\alpha_i} e^{\alpha_i t_i} M^{r-\alpha_i} + r^{\underline{\beta_i+1}} x_i^{\beta_i+1} e^{(\beta_i+1)t_i} M^{r-\beta_i-1} \\
&= \sum_{\alpha_i=0}^{\beta_i+1} S(\beta_i + 1, \alpha_i) r^{\underline{\alpha_i}} x_i^{\alpha_i} e^{\alpha_i t_i} M^{r-\alpha_i},
\end{aligned}$$

which concludes the proof.

□

We now extend the result of Lemma B.3 to an arbitrary derivative of the moment generating function.

**Theorem B.4.** *For any $x \in \Delta_n$, $\beta \in \mathbb{N}^n$ and $t \in \mathbb{R}^n$, one has*

$$\frac{\partial^{|\beta|} M_x^r(t)}{\partial t_1^{\beta_1} \cdots \partial t_n^{\beta_n}} = \sum_{\alpha \in \mathbb{N}^n : \alpha \leq \beta} r^{\underline{|\alpha|}} x^\alpha M_x^{r-|\alpha|}(t) \left( \prod_{i=1}^n e^{\alpha_i t_i} S(\beta_i, \alpha_i) \right).$$

**Proof.** We show the result using induction on the size $k$ of the support of $\beta$, i.e., $k = |\{i \in [n] : \beta_i \neq 0\}|$. The result holds clearly for $k = 0$ and, for $k = 1$, the result holds by Lemma B.3. We now assume that the result holds for $k$ and we show that it also holds for $k+1$. For this, consider the sequences $\beta' = (\beta_1, \ldots, \beta_k, \beta_{k+1}, 0, \ldots, 0)$ and $\beta = (\beta_1, \ldots, \beta_k, 0, 0, \ldots, 0) \in \mathbb{N}^n$, where $\beta_1, \ldots, \beta_{k+1} \geq 1$. By the induction assumption we know that

$$\frac{\partial^{|\beta|} M^r}{\partial t_1^{\beta_1} \cdots \partial t_k^{\beta_k}} = \sum_{0 \leq \alpha \leq \beta} r^{\underline{|\alpha|}} x^\alpha M^{r-|\alpha|} \left( \prod_{i=1}^n e^{\alpha_i t_i} S(\beta_i, \alpha_i) \right), \tag{B.8}$$

setting again $M^r = M_x^r(t)$ for simplicity.
Using (B.8), we obtain

$$\begin{aligned} \frac{\partial^{|\beta'|} M^r}{\partial t_1^{\beta_1} \cdots \partial t_{k+1}^{\beta_{k+1}}} &= \frac{\partial^{\beta_{k+1}}}{\partial t_{k+1}^{\beta_{k+1}}} \frac{\partial^{|\beta|} M^r}{\partial t_1^{\beta_1} \cdots \partial t_k^{\beta_k}} \\ &= \frac{\partial^{\beta_{k+1}}}{\partial t_{k+1}^{\beta_{k+1}}} \left( \sum_{0 \leq \alpha \leq \beta} r^{\underline{|\alpha|}} x^\alpha M^{r-|\alpha|} \left( \prod_{i=1}^n e^{\alpha_i t_i} S(\beta_i, \alpha_i) \right) \right). \end{aligned} \tag{B.9}$$

Note that $\alpha_{k+1} = 0$ since $\alpha_{k+1} \leq \beta_{k+1}$ and $\beta_{k+1} = 0$. Hence, $M^{r-|\alpha|}$ is the only term containing the variable $t_{k+1}$ and thus (B.9) implies

$$\frac{\partial^{|\beta'|} M^r}{\partial t_1^{\beta_1} \cdots \partial t_{k+1}^{\beta_{k+1}}} = \sum_{0 \leq \alpha \leq \beta} r^{\underline{|\alpha|}} x^\alpha \left( \prod_{i=1}^n e^{\alpha_i t_i} S(\beta_i, \alpha_i) \right) \frac{\partial^{\beta_{k+1}} M^{r-|\alpha|}}{\partial t_{k+1}^{\beta_{k+1}}}. \tag{B.10}$$

We now use Lemma B.3 to compute the last term:

$$\frac{\partial^{\beta_{k+1}} M^{r-|\alpha|}}{\partial t_{k+1}^{\beta_{k+1}}} = \sum_{\theta_{k+1}=0}^{\beta_{k+1}} S(\beta_{k+1}, \theta_{k+1})(r - |\alpha|)^{\underline{\theta_{k+1}}} x_{k+1}^{\theta_{k+1}} e^{\theta_{k+1} t_{k+1}} M^{r-|\alpha|-\theta_{k+1}}. \tag{B.11}$$

Substituting (B.11) into (B.10) we obtain

$$
\begin{aligned}
&\frac{\partial^{|\beta'|} M^r}{\partial t_1^{\beta_1} \cdots \partial t_{k+1}^{\beta_{k+1}}} \\
= &\sum_{0 \le \alpha \le \beta} r^{\underline{|\alpha|}} x^\alpha \left( \sum_{\theta_{k+1}=0}^{\beta_{k+1}} S(\beta_{k+1}, \theta_{k+1}) (r - |\alpha|)^{\underline{\theta_{k+1}}} x_{k+1}^{\theta_{k+1}} e^{\theta_{k+1} t_{k+1}} M^{r-|\alpha|-\theta_{k+1}} \right) \times \cdots \\
&\cdots \times \left( \prod_{i=1}^{n} e^{\alpha_i t_i} S(\beta_i, \alpha_i) \right) \\
= &\sum_{0 \le \alpha \le \beta} \sum_{\theta_{k+1}=0}^{\beta_{k+1}} r^{\underline{|\alpha|+\theta_{k+1}}} x^{\alpha + e_{k+1}\theta_{k+1}} S(\beta_{k+1}, \theta_{k+1}) e^{\theta_{k+1} t_{k+1}} M^{r-(|\alpha|+\theta_{k+1})} \times \cdots \\
&\cdots \times \left( \prod_{i=1}^{n} e^{\alpha_i t_i} S(\beta_i, \alpha_i) \right) \\
= &\sum_{0 \le \alpha' \le \beta'} r^{\underline{|\alpha'|}} x^{\alpha'} M^{r-|\alpha'|} \left( \prod_{i=1}^{n} e^{\alpha'_i t_i} S(\beta'_i, \alpha'_i) \right),
\end{aligned}
$$

after setting $\alpha' = \alpha + \mathbf{e}_{k+1}\theta_{k+1}$. This concludes the proof of Theorem B.4. $\square$

Combining (B.3), (B.4) and Theorem B.4, we can conclude the proof for (B.2) and thus Theorem 2.18.

# Appendix C

# Rational minimizers for quadratic optimization

In this appendix, we consider the quadratic optimization problem

$$\begin{aligned} \min \quad & f(x) = x^T H x + c^T x \\ \text{s.t.} \quad & Ax \geq b, \end{aligned} \tag{C.1}$$

where $H \in \mathbb{Z}^{n\times n}, c \in \mathbb{Z}^n, A \in \mathbb{Z}^{t\times n}$ and $b \in \mathbb{Z}^t$. Vavasis [95] shows that there exists a rational global minimizer for problem (C.1). In what follows, we sketch his proof, and then obtain an upper bound for the denominator of the rational global minimizer for the standard quadratic optimization problem.

## C.1 Vavasis' proof

Among all global minimizers of (C.1), let $x^*$ be the one at which the number of active constraints in the system $Ax \geq b$ is maximum. Recall that a constraint is active if it is satisfied with equality. Vavasis [95] shows that $x^*$ is rational. We sketch his approach.

Let $l$ denote the number of entries of $x^*$ that are equal to 0. Let $k$ be the number of active constraints in the system $Ax \leq b$ at $x^*$. Then $x^*$ satisfies the following linear system

$$Mx = b', \tag{C.2}$$

where $M$ denotes the $(k + l) \times n$ matrix, whose first $k$ rows are given by the active constraints in $Ax \geq b$ at $x^*$ and whose last $l$ rows express the constraints that $l$ coordinates of $x^*$ are equal to 0.

Let $r$ denote the rank of $M$. Then we can use elementary row operations to eliminate $r$ variables from $Mx = b'$. That is, one may rewrite $Mx = b'$ as

$$\hat{x} = M'\tilde{x} + b'' \tag{C.3}$$

for some matrix $M' \in \mathbb{Q}^{r\times(n-r)}$ and vector $b'' \in \mathbb{Q}^r$, where $\hat{x} \in \mathbb{R}^r$ and $\tilde{x} \in \mathbb{R}^{n-r}$ give some partition for the vector $x \in \mathbb{R}^n$.

Now denote the feasible set of (C.1) as

$$F := \{x : Ax \geq b\}.$$

We also define

$$F' := \{x \in F : Mx = b' \;\text{ and }\; x_i = 0 \;\text{ if } x_i^* = 0\}. \tag{C.4}$$

If $x \in F'$, then using (C.3), we can express the polynomial $f(x) = x^T Hx + c^T x$ from (C.1) in terms of $\tilde{x}$, say

$$f(x) = \underbrace{\tilde{x}^T \tilde{H}\tilde{x} + \tilde{c}^T\tilde{x}}_{=:\tilde{f}(\tilde{x})} + f_0,$$

for some matrix $\tilde{H} \in \mathbb{Q}^{(n-r)\times(n-r)}$, vector $\tilde{c} \in \mathbb{Q}^{n-r}$ and constant $f_0 \in \mathbb{Q}$.

Now we consider the problem of minimizing $\tilde{f}(\tilde{x})$ on $\mathbb{R}^{n-r}$.

The crucial step in Vavasis' proof is to show that the matrix $\tilde{H}$ is positive definite. Then, the unique global minimizer of $\tilde{f}(\tilde{x})$ must satisfy $\nabla_{\tilde{x}}\tilde{f}(\tilde{x}) = 0$, and it can be calculated from the linear system

$$\tilde{H}\tilde{x} = -\frac{\tilde{c}}{2}. \tag{C.5}$$

This means that $\tilde{f}(\tilde{x})$ has an unique global minimizer in $\mathbb{R}^{n-r}$, namely $\tilde{x} = -\frac{1}{2}\tilde{H}^{-1}\tilde{c}$. Then Vavasis [95] proves that it must be the case that $\tilde{x} = \tilde{x^*}$. Recall that $x^* = (\hat{x^*}, \tilde{x^*})$, where $\hat{x^*}$ is determined from $\tilde{x^*}$ using (C.3).

Thus $x^*$ is determined entirely by solving two systems of equations (C.5) and (C.3). As $\tilde{H}$ and $\tilde{c}$ have rational entries, it follows that $\tilde{x^*}$ is rational (as $\tilde{x^*} = -\frac{1}{2}\tilde{H}^{-1}\tilde{c}$, see relation (C.9) below). Moreover, since the entries of $M'$ and $b''$ are rational, by (C.3), $\hat{x^*}$ is also rational. Hence, the minimizer $x^*$ is rational.

## C.2 Denominator of the rational minimizer

Now we consider the standard quadratic optimization problem:

$$f_{\min,\Delta_n} = \begin{array}{ll} \min & f(x) = x^T Q x \\ \text{s.t.} & x \in \Delta_n, \end{array} \tag{C.6}$$

where $Q \in \mathbb{Z}^{n\times n}$. The standard simplex can be written as

$$\Delta_n = \left\{ x \in \mathbb{R}^n \ : \ \sum_{i=1}^n x_i \geq 1, -\sum_{i=1}^n x_i \geq -1, x_i \geq 0 \ (i = 1, \ldots, n) \right\}.$$

As we saw above in Section C.1, by putting (C.6) in the form of (C.1), problem (C.6) has a rational minimizer $x^*$.

Next, we show an upper bound for the denominator of $x^*$.

**Lemma C.1.** *Set $Q_{\max} := \max_{i,j} |Q_{ij}|$. Let $x^* \in \Delta(n,m)$ be a rational minimizer of problem* (C.6)*, obtained as explained in Section C.1. Assume $l$ coordinates of $x^*$ are zeros. Then, the denominator of $x^*$ can be upper bounded as*

$$m \leq (4Q_{\max})^{n-l-1}.$$

In the rest of this subsection, we give the proof of Lemma C.1.

By putting program (C.6) in the form of program (C.1), we have $H = Q$, $c = 0$,

$$A = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ -1 & -1 & \cdots & -1 \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \ddots & \cdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \quad \text{and} \quad b = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

First, we analyze the linear systems (C.2) and (C.5), and then we give an upper bound for $m$.

**Step 1: analysis of the system (C.2)**

First, we assume w.l.o.g. that the first $l$ coordinates of $x^*$ are zeroes. Then, define the matrix $M$ appearing in (C.2):

$$M = \begin{bmatrix} \mathbf{e}^T & \mathbf{e}^T \\ -\mathbf{e}^T & -\mathbf{e}^T \\ I_l & \mathbf{0} \\ I_l & \mathbf{0} \end{bmatrix},$$

where $I_l \in \mathbb{R}^{l\times l}$ denotes the identity matrix and $\mathbf{0}$ denotes the zero matrix of suitable size. In this case, we have $k = l + 2$ and $M$ has rank $r = l + 1$. Then we have

$$\hat{x^*} = \begin{bmatrix} \mathbf{0} \\ 1 - \sum_{i=r+1}^{n} x_i^* \end{bmatrix} \in \mathbb{R}^r, \quad \text{i.e.,} \quad \hat{x^*} = \begin{bmatrix} \mathbf{0} \\ -\mathbf{e}^T \end{bmatrix} \tilde{x^*} + \mathbf{e}_r, \tag{C.7}$$

where $\mathbf{e}_r \in \mathbb{R}^r$ denotes the $r$-th unit vector.

**Step 2: analysis of the system (C.5)**

When $x \in F'$ (in (C.4)), in order to use (C.5), we need the matrix $\tilde{H}$ and the vector $\tilde{c}$ such that

$$f(x) = x^T Q x = \tilde{x}^T \tilde{H} \tilde{x} + \tilde{c}^T \tilde{x} + f_0. \tag{C.8}$$

**Lemma C.2.** *Let $\tilde{H}$ and $\tilde{c}$ in* (C.8) *be indexed by $\{r+1, r+2, \ldots, n\}$. Then, for any $i, j \in \{r+1, r+2, \ldots, n\}$, one has*

$$\begin{aligned} \tilde{H}_{ij} &= Q_{ij} + Q_{rr} - Q_{ri} - Q_{rj}, \\ \tilde{c}_i &= 2Q_{ri} - 2Q_{rr}, \\ f_0 &= Q_{rr}. \end{aligned}$$

**Proof.** Since

$$\hat{x} = \begin{bmatrix} \mathbf{0} \\ 1 - \sum_{i=r+1}^{n} x_i \end{bmatrix} \in \mathbb{R}^r,$$

one has

$$\begin{aligned}
f(x) &= x^T Q x \\
&= 2 \sum_{i,j: r \le i < j \le n} Q_{ij} x_i x_j + \sum_{i=r}^{n} Q_{ii} {x_i}^2 \\
&= \sum_{i=r+1}^{n} Q_{ii} {x_i}^2 + Q_{rr} (1 - \sum_{i=r+1}^{n} x_i)^2 + 2 \Bigg[ \sum_{i,j: r+1 \le i < j \le n} Q_{ij} x_i x_j \\
&\quad + \sum_{j: r+1 \le j \le n} Q_{rj} x_j (1 - \sum_{i=r+1}^{n} x_i) \Bigg] \\
&= \sum_{i=r+1}^{n} {x_i}^2 (Q_{ii} + Q_{rr} - 2Q_{ri}) + \sum_{i,j: r+1 \le i < j \le n} x_i x_j (2Q_{ij} + 2Q_{rr} - 2Q_{ri} - 2Q_{rj}) \\
&\quad + \sum_{i=r+1}^{n} x_i (2Q_{ri} - 2Q_{rr}) + Q_{rr}.
\end{aligned}$$

This finishes the proof. □

**Step 3: an upper bound for the denominator $m$**

Note that the minimizer $x^*$ is of the form $(\hat{x^*}, \tilde{x^*})$ and $\hat{x^*}$ is determined by $\tilde{x^*}$ from (C.7). Then, recall that $\tilde{x^*}$ is determined by the linear system (C.5), for which the entries of the matrix $\tilde{H}$ and the vector $\tilde{c}$ are given in Lemma C.2. By Cramer's rule, the $i$-th entry of $\tilde{x^*}$ is given by

$$\tilde{x^*_i} = \frac{\det \tilde{H}_i}{\det \tilde{H}}, \quad \text{for any } i \in [n-r], \tag{C.9}$$

where $\tilde{H}_i$ is the matrix formed by replacing the $i$-th column of $\tilde{H}$ by the column vector $-\tilde{c}/2$.

By Lemma C.2, both $\det \tilde{H}_i$ and $\det \tilde{H}$ are integral. Thus, the denominator of $\tilde{x^*}$ can be upper bounded by $|\det \tilde{H}|$, which is equal to $\det \tilde{H}$ since $\tilde{H}$ is positive definite.

From (C.7), we know that the entries of $\hat{x^*}$ are zeroes except the last one which is equal to $1 - \sum_{i=1}^{n-r} \tilde{x^*}_i$. Together with the fact that all entries of $\tilde{x^*}$ have a common denominator $\det \tilde{H}$ (see (C.9)), we can upper bound the denominator of $x^*$ by $\det \tilde{H}$, i.e.,

$$m \le \det \tilde{H}.$$

Since $\tilde{H} \succ 0$ one may use the Hadamard inequality to bound its determinant:

$$\det \tilde{H} \leq \prod_i \tilde{H}_{ii},$$

see, e.g., [39, Theorem 7.8.1]. Recall that we denote the largest entry (in absolute value) of $Q$ as $Q_{\max}$. Then, by Lemma C.2, any diagonal entry of $\tilde{H}$ is at most $4Q_{\max}$. Thus:

$$m \leq (4Q_{\max})^{n-r},$$

where $r = l + 1$ and $l$ coordinates of $x^*$ are zeros.

This finishes the proof of Lemma C.1.

# Bibliography


[1] Abramowitz, M., Stegun, I.A.: *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables.* National Bureau of Standards Applied Mathematics Series 55 (1972)

[2] Altomare, F., Campiti, M.: *Korovkin-type Approximation Theory and Its Applications.* De Guyter studies in mathematics 17, Walter de Guyter publishers, Berlin (1994)

[3] Anjos, M., Lasserre, J.B.: *Handbook of Semidefinite, Conic and Polynomial Optimization.* Springer, New York (2011)

[4] Barahona, F., Mahjoub, A.R.: On the cut polytope. *Math. Program.* 36, 157–173 (1986)

[5] Bellare, M., Rogaway, P.: The complexity of approximating a nonlinear program. *Math. Program.* 69, 429–441 (1995)

[6] Bland, R.G., Goldfarb, D, Todd, M.J.: The ellipsoid method: A survey. *Oper. Res.* 29(6), 1039–1091 (1981)

[7] Blumenson, L.E.: A derivation of $n$-dimensional spherical coordinates. *Amer. Math. Monthly.* 67(1), 63–66 (1960)

[8] Bomze, I.M., Klerk, E. de.: Solving standard quadratic optimization problems via semidefinite and copositive programming. *J. Global Optim.* 24(2), 163–185 (2002)

[9] Bomze, I.M., Gollowitzer, S., Yildirim, E.A.: Rounding on the standard simplex: Regular grids for global optimization. *J. Global Optim.* 59(2–3), 243–258 (2014)

[10] Bos, L.P.: Bounding the Lebesque function for Lagrange interpolation in a simplex. *J. Approx. Theory.* 38, 43–59 (1983)

[11] Bruck, J., Blaum, M.: Neural networks, error-correcting codes, and polynomials over the binary $n$-cube. *IEEE Trans. Inform. Theory.* 35(5), 976–987 (1989)

[12] Chudnovsky, M., Robertson, N., Seymour, P., Thomas, R.: The strong perfect graph theorem. *Ann. Math.* 164(1), 51–229 (2006)

[13] Cornaz, D., Jost, V.: A one-to-one correspondance between colorings and stable sets. *Oper. Res. Lett.* 36(6), 673–676 (2008)

[14] Faybusovich, L.: Global optimization of homogeneous polynomials on the simplex and on the sphere. In: Floudas, C., Pardalos, P. (eds.) *Frontiers in Global Optimization,* pp. 109–121. Kluwer Academic Publishers, Dordrecht (2003)

[15] de Klerk, E.: *Aspects of Semidefinite Programming: Interior Point Algorithms and Selected Applications.* In series: Applied Optimization, Vol 65, Springer (2002)

[16] de Klerk, E.: The complexity of optimizing over a simplex, hypercube or sphere: a short survey. *Cent. Eur. J. Oper. Res.* 16(2), 111–125 (2008)

[17] de Klerk, E., den Hertog, D, Elabwabi, G.: On the complexity of optimization over the standard simplex. *European J. Oper. Res.* 191(3), 773–785 (2008)

[18] de Klerk, E., Laurent, M.: Error bounds for some semidefinite programming approaches to polynomial optimization on the hypercube. *SIAM J. Optim.* 20(6), 3104–3120 (2010)

[19] de Klerk, E., Laurent, M.: On the Lasserre hierarchy of semidefinite programming relaxations of convex polynomial optimization problems. *SIAM J. Optim.* 21, 824-832 (2011)

[20] de Klerk, E., Laurent, M., Parrilo, P.: On the equivalence of algebraic approaches to the minimization of forms on the simplex. In: Henrion, D.,Garulli, A. (eds.) *Positive Polynomials in Control,* pp. 121–133. Springer Verlag, Berlin (2005)

[21] de Klerk, E., Laurent, M., Parrilo, P.: A PTAS for the minimization of polynomials of fixed degree over the simplex. *Theoret. Comput. Sci.* 361(2–3), 210–225 (2006)

[22] de Klerk, E., Laurent, M., Sun, Z.: An alternative proof of a PTAS for fixed-degree polynomial optimization over the simplex. *Math. Program.* DOI: 10.1007/s10107-014-0825-6 (2014)

[23] de Klerk, E., Laurent, M., Sun, Z.: An error analysis for polynomial optimization over the simplex based on the multivariate hypergeometric distribution. *SIAM J. Optim.* (accepted with minor revision) (2015)

[24] de Klerk, E., Laurent, M., Sun, Z.: Convergence analysis for Lasserre's measure-based hierarchy of upper bounds for polynomial optimization. Preprint at arXiv: 1411.6867 (2014)

[25] de Klerk, E., Pasechnik, D.V.: Approximating of the stability number of a graph via copositive programming. *SIAM J. Optim.* 12(4), 875–892 (2002)

[26] de Loera, J., Lee, J., Margulies, S., Onn, S.: Expressing combinatorial problems by systems of polynomial equations and Hilbert's Nullstellensatz. *Combin. Probab. Comput.* 18(4), 551–582 (2009)

[27] de Loera, J., Rambau, J., Santos, F.: *Triangulations: Structures for Algorithms and Applications.* In series: Algorithms and Computation in Mathematics, Vol. 25, Springer (2010)

[28] Ditzian, Z.: Inverse theorems for multidimensional Bernstein operators. *Pacific J. Math.* 121 (2), 293–319 (1986)

[29] Ditzian, Z.: Best polynomial approximation and Bernstein polynomial approximation on a simplex, *Indag. Math. (N.S.).* 92, 243–256 (1989)

[30] Ditzian, Z., Zhou, X.: Optimal approximation class for multivariate Bernstein operators, *Pacific J. Math.* 158, 93–120 (1993)

[31] Doherty, A.C., Wehner, S.: Convergence of SDP hierarchies for polynomial optimization on the hypersphere. arXiv:1210.5048v2 (2013)

[32] Dyer, M.E., Frieze, A.M.: On the complexity of computing the volume of a polyhedron. *SIAM J. Comput.* 17(5), 967–974 (1988)

[33] Faybusovich, L.: Global optimization of homogeneous polynomials on the simplex and on the sphere. In C. Floudas and P. Pardalos, editors, Frontiers in Global Optimization. Kluwer Academic Publishers (2003)

[34] Gouveia, J., Parrilo, P., Thomas, R.: Theta bodies for polynomial ideals. *SIAM J. Optim.* 20(4), 2097–2118 (2010)

[35] Grant, M., Boyd, S.: Graph implementations for nonsmooth convex programs. Recent Advances in Learning and Control (tribute to M. Vidyasagar), V. Blondel, S. Boyd, and H. Kimura, editors, pages 95–110, *Lecture Notes in Control and Information Sciences*, Springer (2008)

[36] Grant, M., Boyd, S.: CVX: Matlab software for disciplined convex programming, version 2.0 beta. http://cvxr.com/cvx (2013).

[37] Grundmann, A., Moeller, H.M.: Invariant integration formulas for the n-simplex by combinatorial methods. *SIAM J. Numer. Anal.* 15, 282–290 (1978)

[38] Handelman, D.: Representing polynomials by positive linear functions on compact convex polyhedra. *Pacific J. Math.* 132(1), 35–62 (1988)

[39] Horn, R.A., Johnson, C.R.: *Matrix Analysis.* Cambridge University Press (1990)

[40] Johnson, N.L., Kotz, S., Balakrishnan, N.: *Discrete Multivariate Distributions.* Wiley (1997)

[41] Karp, R.M.: Reducibility among combinatorial problems. In: Miller, R.E., Thatcher, J.W. (eds.) *Complexity of Computer Computations*, 85–103. Springer, New York (1972)

[42] Kleniati, P.M., Rustem, B.: Portfolio decisions with higher order moments. *COMISEF Working Papers Series,* WPS-021 (2009)

[43] Kleniati, P.M., Parpas, P., Rustem, B.: Partitioning procedure for polynomial optimization: application to portfolio decisions with higher order moments. *COMISEF Working Papers Series,* WPS-023 (2009)

[44] Knoblauch, A.: Closed-form expressions for the moments of the binomial probability distribution. *SIAM J. Appl. Math.* 69(1), 197–204 (2008)

[45] Krivine, J.L.: Anneaux préordonnés, *J. Anal. Math.* 12, 307–326 (1964)

[46] Krivine, J.L.: Quelques propriétés des préordres dans les anneaux commutatifs unitaires. *Comptes Rendus de l'Académie des Sciences de Paris*, 258, 3417–3418 (1964)

[47] Lasserre, J.B.: Global optimization with polynomials and the problem of moments. *SIAM J. Optim.* 11, 796–817 (2001)

[48] Lasserre, J.B., Zeron, E.S.: Solving a class of multivariate integration problems via Laplace techniques. *Appl. Math. (Warsaw).* 28(4), 391–405 (2001)

[49] Lasserre, J.B.: An explicit equivalent positive semidefinite program for nonlinear 0-1 programs. *SIAM J. Optim.* 12, 756–769 (2002)

[50] Lasserre, J.B.: Semidefinite programming vs. LP relaxations for polynomial programming. *Math. Oper. Res.* 27(2), 347–360 (2002)

[51] Lasserre, J.B.: Convexity in semialgebraic geometry and polynomial optimization. *SIAM J. Optim.* 19, 1995–2014 (2009)

[52] Lasserre, J.B.: *Moments, Positive Polynomials and Their Applications.* Imperial College Press (2009)

[53] Lasserre, J.B.: A new look at nonnegativity on closed sets and polynomial optimization. *SIAM J. Optim.* 21, 864–885 (2011)

[54] Lasserre, J.B.: Unit balls of constant volume: which one has optimal representation? Preprint at arXiv: 1408.1324 (2014)

[55] Laurent, M.: A comparison of the Sherali-Adams, Lovász-Schrijver and Lasserre relaxation for 0-1 programming. *Math. Oper. Res.* 28(3), 470–498 (2003)

[56] Laurent, M., Rendl, F.: Semidefinite Programming and Integer Programming. In: Aardal, K., Nemhauser, G., Weismantel, R. (Eds.) *Handbook on Discrete Optimization*, Vol 12, 393–514, Elsevier B.V. (2005)

[57] Laurent, M.: Semidefinite representations for finite varieties. *Math. Program.* 109, 1–26 (2007)

[58] Laurent, M.: Sums of squares, moment matrices and optimization over polynomials. In: Putinar, M., Sullivant, S. (eds.) *Emerging Applications of Algebraic Geometry*, pp. 157–270. Springer, New York (2009)

[59] Laurent, M., Sun, Z.: Handelman's hierarchy for the maximum stable set problem. *J. Global Optim.* 60(3), 393–423 (2014)

[60] Law, A.M.: *Simulation Modeling and Analysis* (4th edition). Mc Graw-Hill (2007)

[61] Li, Z., He, S., Zhang, S.: *Approximation Methods for Polynomial Optimization: Models, Algorithms and Applications.* Springer (2012)

[62] Lipták, L., Tuncel, L.: The stable set problem and the lift-and-project ranks of graphs. *Math. Program.* 98(1-3), 319–353 (2003)

[63] Lovász, L.: A characterization of perfect graphs. *J. Combin. Theory Ser. B.* 13(2), 95–98 (1972)

[64] Lovász, L.: Stable sets and polynomials. *Discrete Math.* 124(1-3), 137–153 (1994)

[65] Lovász, L., Schrijver, A.: Cones of matrices and set-functions and $0-1$ optimization. *SIAM J. Optim.* 1(2), 166–190 (1991)

[66] Markowitz, H.M.: Portfolio selection. *J. Finance.* 7, 79–91 (1952)

[67] Marshall, M.: *Positive Polynomials and Sums of Squares.* Mathematical Surveys and Monographs, vol. 146 (2008)

[68] Mattila, P.: *Geometry of Sets and Measures in Euclidean Spaces: Fractals and Rectifiability.* Cambridge University Press (1999)

[69] Motwani, R., Raghavan, P.: *Randomized Algorithms.* Cambridge University Press (1995)

[70] Motzkin, T.S., Straus, E.G.: Maxima for graphs and a new proof of a theorem of Túran, *Canad. J. Math.* 17, 533–540 (1965)

[71] Nemhauser, G.L., Trotter Jr, L.E.: Properties of vertex packing and independence system polyhedra. *Math. Program.* 6(1), 48–61 (1974)

[72] Nesterov, Y.: Random walk in a simplex and quadratic optimization over convex polytopes. *CORE Discussion Paper* 2003/71, CORE-UCL, Louvain-La-Neuve (2003)

[73] Nesterov, Y., Wolkowicz, H., Ye, Y.: Semidefinite programming relaxations of nonconvex quadratic optimization. In: H. Wolkowicz, R. Saigal, L. Vandenberghe (eds.), *Handbook of Semidefinite Programming*, Kluwer Academic Publishers, Norwell, MA, 361–419 (2000)

[74] Nie, J., Schweighofer, M.: On the complexity of Putinar's Positivstellensatz. *J. Complexity.* 23(1), 135–150 (2007)

[75] Nie, J.: Certifying convergence of Lasserre's hierarchy via flat truncation. *Math. Program.* 142(1–2), 485–510 (2013)

[76] Nie, J.: Polynomial optimization with real varieties. *SIAM J. Optim.* 23(3), 1634–1646 (2013)

[77] Park, M.-J., Hong, S.-P.: Rank of Handelman hierarchy for Max-Cut. *Oper. Res. Lett.* 39(5), 323–328 (2011)

[78] Park, M.-J., Hong, S.-P.: Handelman rank of zero-diagonal quadratic programs over a hypercube and its applications. *J. Global Optim.* 56(2), 727–736 (2013)

[79] Parrilo, P.: Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization. Ph.D. Thesis, California Institute of Technology, May 2000.

[80] Parrilo, P.: Semidefinite programming relaxations for semialgebraic problems. *Math. Program.* 96, 293–320 (2003)

[81] Peña, J.C., Vera, J.C., Zuluaga, L.F.: Computing the stability number of a graph via linear and semidefinite programming. *SIAM J. Optim.* 18(1), 87–105 (2007)

[82] Pólya, G.: Collected Papers, Vol.2, MIT Press, Cambridge, MA, London, 309–313 (1974)

[83] Powers, V., Reznick, B.: A new bound for Pólya's theorem with applications to polynomials positive on polyhedra. *J. Pure Appl. Algebra.* 164, 221–229 (2001)

[84] Putinar, M.: Positive polynomials on compact sem-algebraic sets. *Indiana Univ. Math. J.* 42, 969–984 (1993)

[85] Renegar, J.: *A Mathematical View of Interior-Point Methods for Convex Optimization.* MPS-SIAM Series on Optimization, Philadelphia (2001).

[86] Sagol, G., Yildirim, E.A.: Analysis of copositive optimization based linear programming bounds on standard quadratic optimization. *J. Global Optim.* DOI: 10.1007/s10898-015-0269-4 (2015)

[87] Schmüdgen, K.: The K-moment problem for compact semi-algebraic sets. *Math. Ann.* 289, 203–206 (1991)

[88] Schrijver, A.: *Combinatorial Optimization - Polyhedra and Efficiency.* Springer-Verlag, Berlin (2003)

[89] Schweighofer, M.: On the complexity of Schmüdgen's Positivstellensatz. *J. Complexity.* 20, 529–543 (2004)

[90] Sherali, H.D., Adams, W.P.: A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems. *SIAM J. Discrete Math.* 3(3), 411–430 (1990)

[91] Shor, N.Z.: The development of numerical methods for nonsmooth optimization in the USSR. In J.K. Lenstra, A.H.G. Rinnooy Kan, and A. Schrijver, (eds): *History of Mathematical Programming: A Collection of Personal Reminiscences*, pp 135–139, Centrum voor Wiskunde en Informatica and North-Holland, Amsterdam (1991)

[92] Sun, Z.: A refined error analysis for fixed-degree polynomial optimization over the simplex. *J. Oper. Res. Soc. China.* 2(3), 379–393 (2014)

[93] Toh, K.C., Todd, M.J., Tutuncu, R.H.: SDPT3-a Matlab software package for semidefinite programming. *Optim. Methods Softw.* 11, 545–581 (1999)

[94] Tutuncu, R.H., Toh, K.C., Todd, M.J.: Solving semidefinite-quadratic-linear programs using SDPT3. *Math. Program.* 95, 189–217 (2003)

[95] Vavasis, S.A.: Quadratic programming is in NP. *Inform. Process. Lett.* 36, 73–77 (1990)

[96] Vavasis, S.A.: Approximation algorithms for concave quadratic programming. In: C.A. Floudas, P. Pardalos (eds.), *Recent Advances in Global Optimization*, Princeton University Press, Princeton, NJ, 3–18 (1998)

[97] Wendland, H.: *Scattered Data Approximation.* Cambridge University Press (2005)

[98] Test functions for optimization. In *Wikipedia.* Retrived Feburary 2015, from `http://en.wikipedia.org/wiki/Test_functions_for_optimization`

[99] Whittaker, E.T., Watson, G.W.: *A Course of Modern Analysis* (4ed). Cambridge University Press, New York (1996)

[100] Yildirim, E.A.: On the accuracy of uniform polyhedral approximations of the copositive cone. *Optim. Methods Softw.* 27(1), 155–173 (2012)

# List of Symbols

## Sets

| | |
|---|---|
| $\mathbb{R}$ | The set of real numbers. |
| $\mathbb{R}_+$ | The set of nonnegative real numbers. |
| $\mathbb{Q}$ | The set of rational numbers. |
| $\mathbb{Z}$ | The set of integers. |
| $\mathbb{N}$ | The set of nonnegative integers. |
| $\mathbb{R}^n$ | The set of $n$-dimensional real vectors. |
| $\mathbb{R}^n_+$ | The set of $n$-dimensional nonnegative real vectors. |
| $\mathbb{Q}^n$ | The set of $n$-dimensional rational vectors. |
| $\mathbb{Z}^n$ | The set of $n$-dimensional integer vectors. |
| $\mathbb{N}^n$ | The set of $n$-dimensional nonnegative integer vectors. |
| $\mathbb{R}[x]$ | The set of polynomials with real coefficients. |
| $\mathbb{R}_+[x]$ | The set of polynomials with nonnegative real coefficients. |
| $\mathbb{R}[x]_r$ | The set of polynomials with degree at most $r$. |
| $\mathbb{R}_+[x]_r$ | The set of polynomials with nonnegative real coefficients and with degree at most $r$. |
| $\Sigma[x]$ | The set of sums of squares of polynomials. |
| $\Sigma[x]_r$ | The set of sums of squares of polynomials with degree at most $2r$. |
| $\mathcal{H}_{n,d}$ | The set of $n$-variate homogeneous polynomials with degree $d$. |
| $\mathscr{P}(\mathbf{K})$ | The set of polynomials that are nonnegative on $\mathbf{K}$. |
| $[n]$ | This is defined as $\{1, 2, \ldots, n\}$. |
| $\mathbf{e}$ | The all-ones vector. |
| $\mathbf{e}_i$ | The $i$-th standard unit vector. |
| $\mathbf{e}_I$ | This is defined as $\sum_{i \in I} \mathbf{e}_i$. |
| $\mathbf{Q}_n$ | The unit hypercube $[0, 1]^n$. |

| | |
|---|---|
| $\mathbf{B}_\epsilon(a)$ | The Euclidean ball with center $a$ and radius $\epsilon$. |
| $\Delta_n$ | The standard simplex $\{x \in \mathbb{R}^n_+ : \sum_{i=1}^n x_i = 1\}$. |
| $\widehat{\Delta_n}$ | The full-dimensional simplex $\{x \in \mathbb{R}^n_+ : \sum_{i=1}^n x_i \leq 1\}$. |
| $I(n,r)$ | This is defined as $\{x \in \mathbb{N}^n : \sum_{i=1}^n x_i = r\}$. |
| $\Delta(n,r)$ | This is defined as $\{x \in \Delta_n : rx \in \mathbb{N}^n\}$. |
| $N(n,r)$ | This is defined as $\{x \in \mathbb{N}^n : \sum_{i=1}^n x_i \leq r\}$. |
| $\mathcal{P}_t(V)$ | This is defined as $\{I \subseteq V : \vert I\vert \leq t\}$. |
| $\mathcal{P}_{=t}(V)$ | This is defined as $\{I \subseteq V : \vert I\vert = t\}$. |
| $\mathcal{S}_n$ | The set of symmetric $n \times n$ matrices. |
| $\mathcal{S}_n^+$ | The set of $n \times n$ positive semidefinite matrices. |
| $\mathcal{C}_n$ | The set of $n \times n$ copositive matrices. |

# Polynomials and Functions

| | |
|---|---|
| $B_r(f)(x)$ | The Bernstein approximation of order $r$ of a continuous function $f$ on the simplex. |
| $x^\alpha$ | This is defined as $x_1^{\alpha_1} \cdots x_n^{\alpha_n}$. |
| $\vert\alpha\vert$ | This is defined as $\sum_{i=1}^n \alpha_i$. |
| $\phi_\alpha(x)$ | This is defined as $x^\alpha$. |
| $x^{\underline{d}}$ | This is defined as $x(x-1)(x-2)\cdots(x-d+1)$. |
| $x^{\underline{\alpha}}$ | This is defined as $\prod_{i=1}^n x_i^{\underline{\alpha_i}}$. |
| $\alpha!$ | This is defined as $\alpha_1!\alpha_2!\cdots\alpha_n!$. |
| $\Gamma(\cdot)$ | Euler gamma function. |
| $\Vert x\Vert$ | $\ell_2$-norm of $x$. |
| $m_\alpha(\mathbf{K})$ | The moment $\int_{\mathbf{K}} x^\alpha dx$. |
| $w_{\min}(\mathbf{K})$ | The minimum distance between two distinct parallel supporting hyperplanes of $\mathbf{K}$. |
| $D(\mathbf{K})$ | This is defined as $\sup_{x,y\in\mathbf{K}} \Vert x-y\Vert^2$. |
| $S(n,k)$ | Stirling number of the second kind. |
| $k!!$ | Double factorial of $k$. |
| $O(\cdot)$ | The big-oh notation. |
| $\Omega(\cdot)$ | The big-omega notation. |

# Graphs

| | |
|---|---|
| $\overline{G}$ | The complementary graph of graph $G$. |

| | |
|---|---|
| $K_n$ | The complete graph on $n$ vertices. |
| $C_n$ | The circuit on $n$ vertices. |
| $\alpha(G)$ | The stability number of graph $G$. |
| $\omega(G)$ | The clique number of graph $G$. |
| $\chi(G)$ | The chromatic number of graph $G$. |

# Index